Multivariate General Linear Models (MGLM) on Riemannian Manifolds

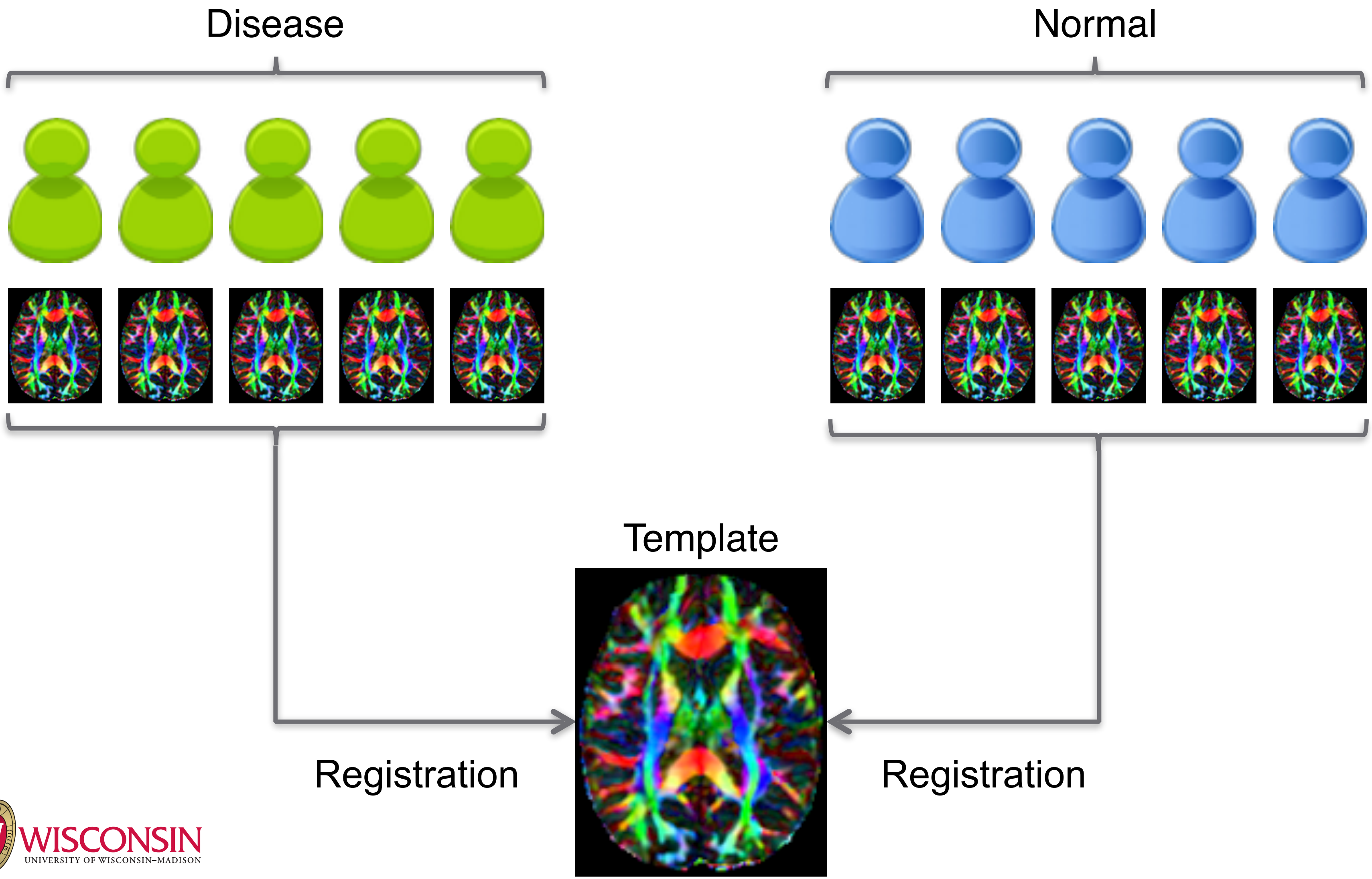
Hyunwoo J. Kim, Nagesh Adluru, Maxwell D. Collins, Moo K. Chung,
Barbara B. Bendlin, Sterling C. Johnson, Richard J. Davidson, Vikas Singh

http://pages.cs.wisc.edu/~hwkim/projects/riem-mglm/

Poster session ID: O-3C-384

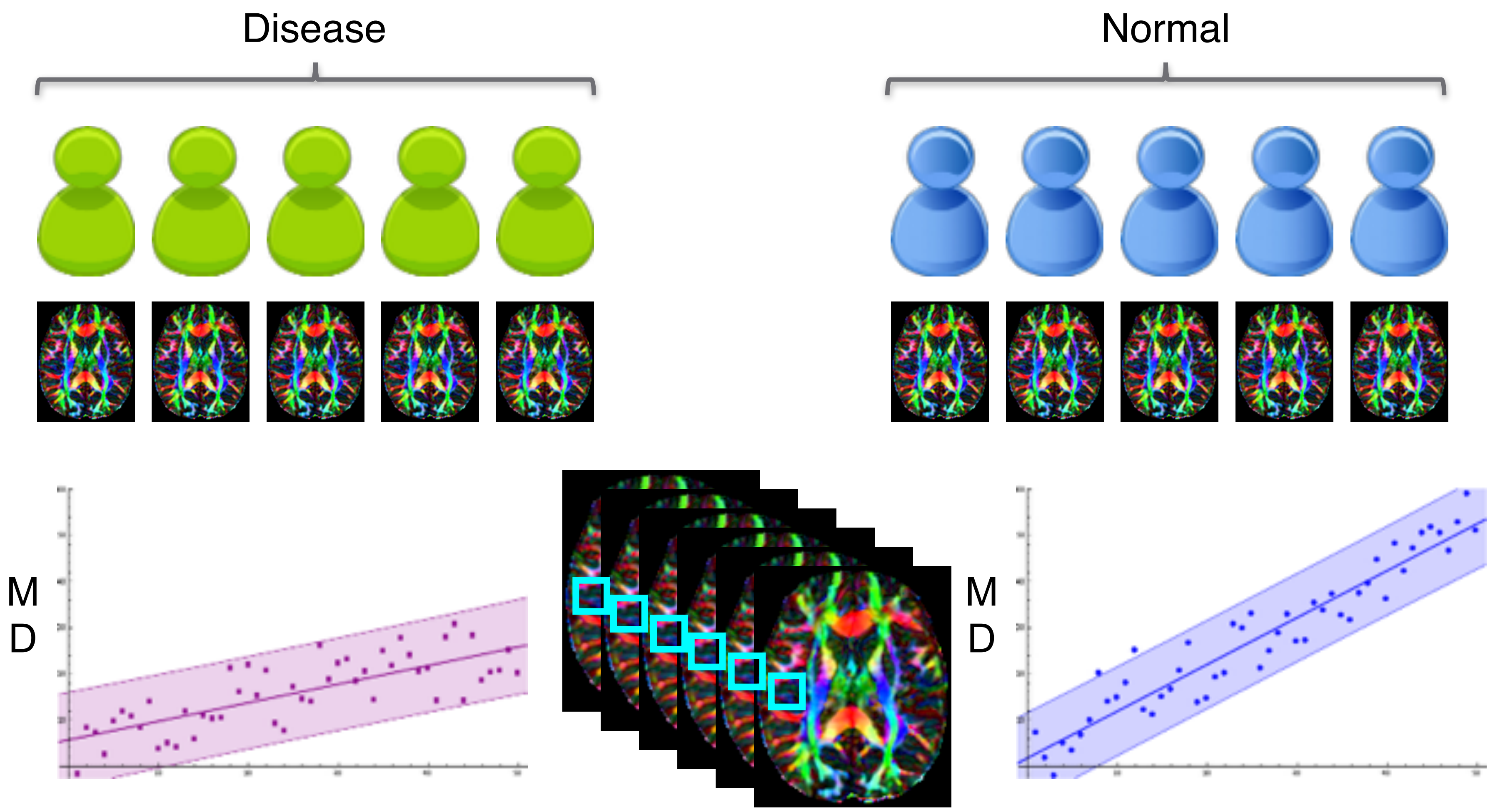
WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

General Linear Model for Group Analysis



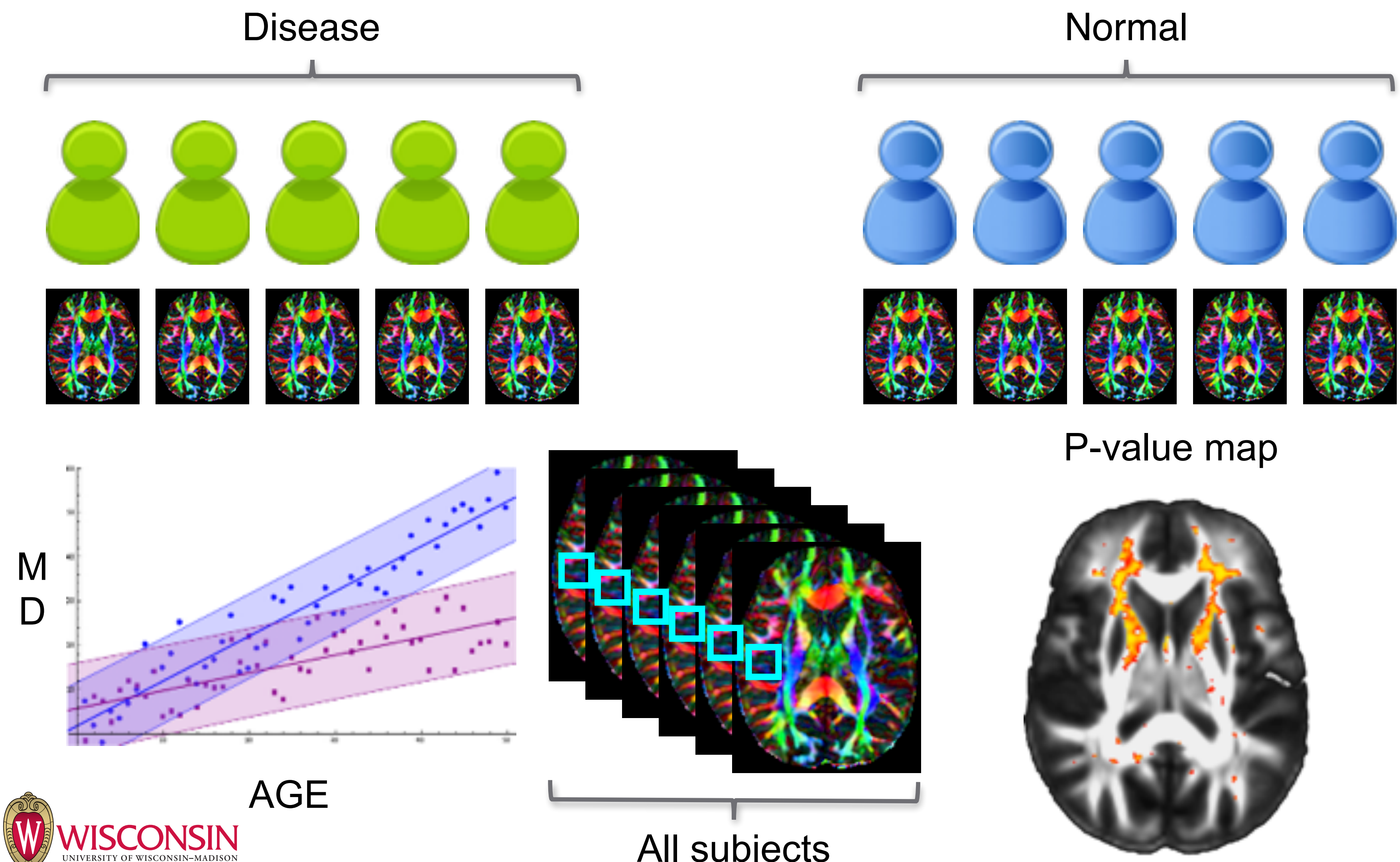
WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

General Linear Model for Group Analysis



WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

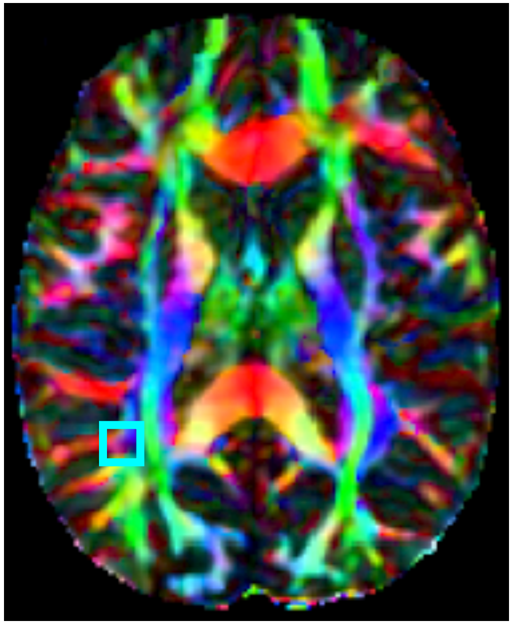
General Linear Model for Group Analysis



WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

General Linear Model (GLM)

$$f : R \to R$$



$$y = \alpha + \beta^1 x^1 + \epsilon$$

x^1: Group (patient or normal)

General Linear Model (GLM)

$$f : R^2 \rightarrow R$$

$$y = \alpha + \beta^1 x^1 + \beta^2 x^2 + \epsilon$$

x^1: Group (patient or normal)

x^2: Age

General Linear Model (GLM)

$$f : R^3 \to R$$

$$y = \alpha + \beta^1 x^1 + \beta^2 x^2 + \beta^3 x^3 + \epsilon$$

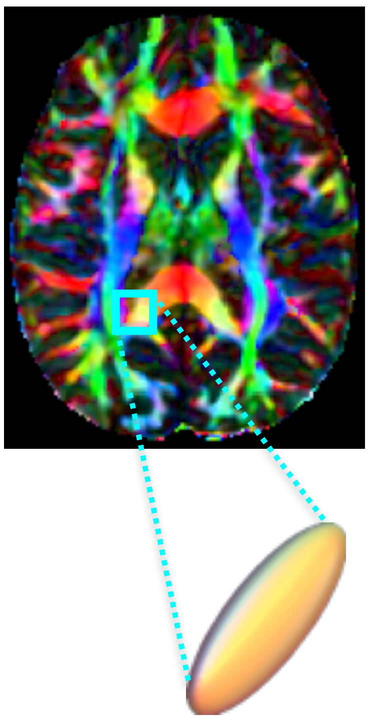
x^1: Group (patient or normal)

x^2: Age

x^3: Gender

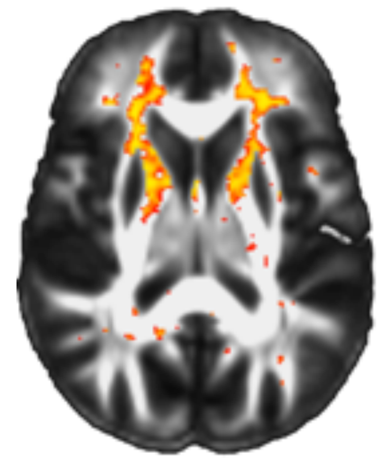
GLM on scalar valued summaries

DTI

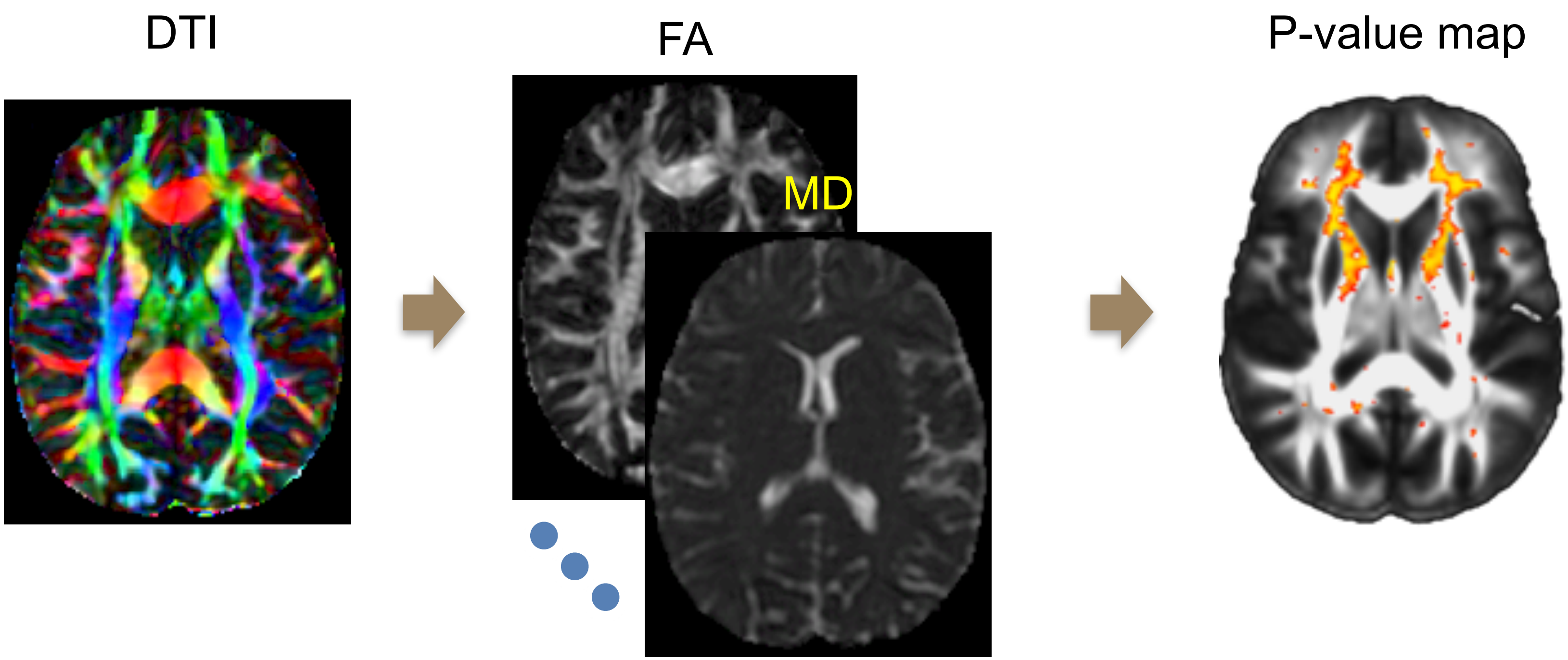


$$D = \begin{pmatrix} 1.53 & 1.38 & 0.65 \\ 1.38 & 1.33 & 0.70 \\ 0.65 & 0.70 & 1.06 \end{pmatrix}$$

P-value map



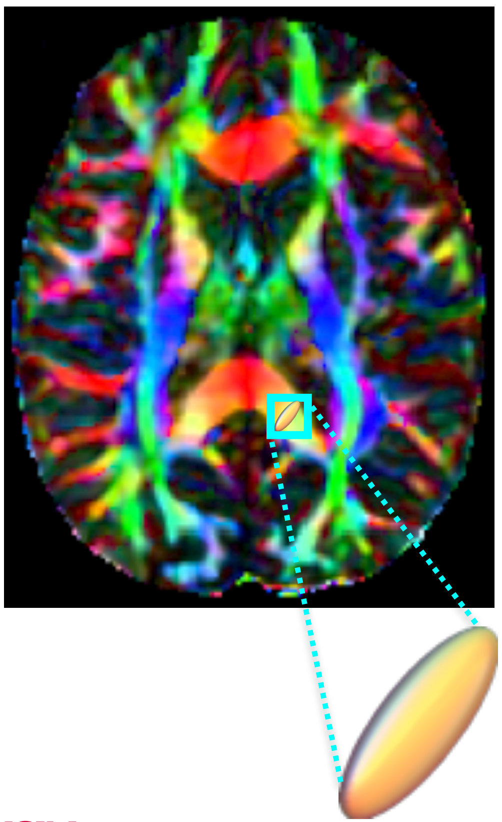
GLM on scalar valued summaries



WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

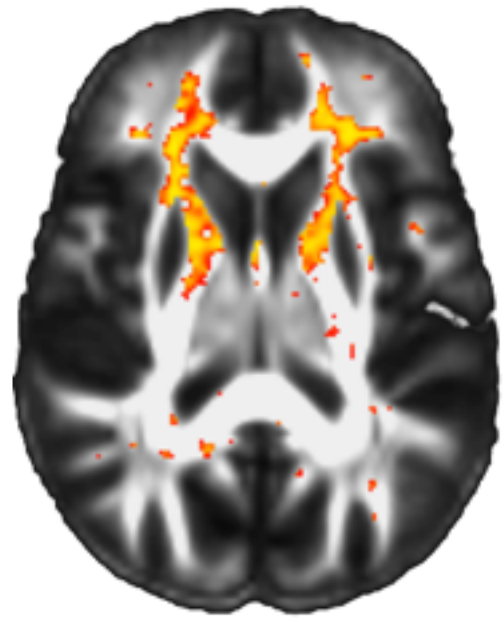
GLM on scalar valued summaries

DTI



?

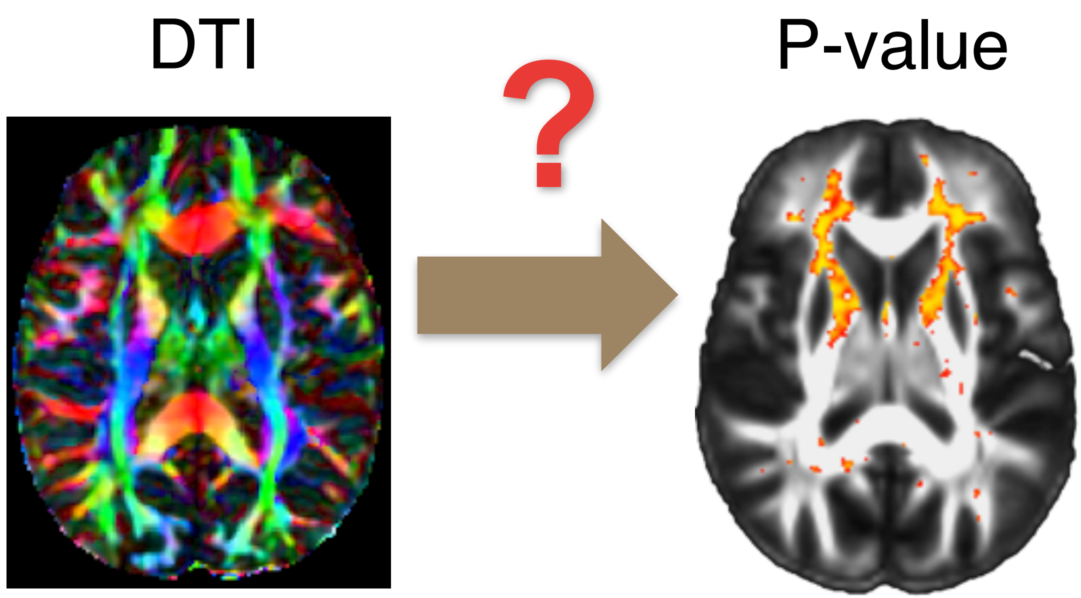
P-value map



$$D = \begin{pmatrix} 1.53 & 1.38 & 0.65 \\ 1.38 & 1.33 & 0.70 \\ 0.65 & 0.70 & 1.06 \end{pmatrix} \quad x^T D x > 0, x \neq 0$$

WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

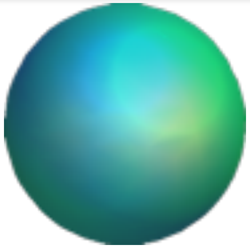
GLM on scalar valued summaries



Response Y is manifold-valued

Manifold-valued data

- Unit sphere, quotient spaces of spheres



- SPD matrix i.e covariance matrix, diffusion tensors

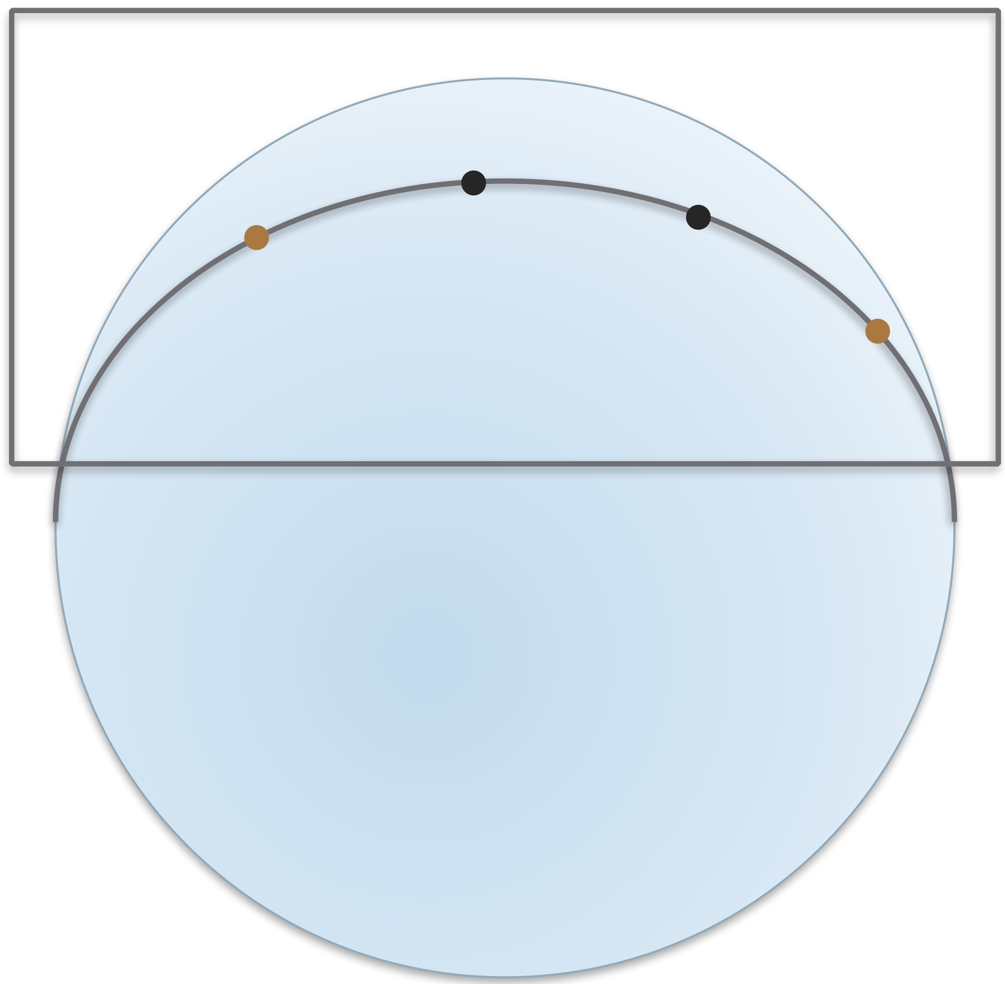
- Probability density functions (PDFs), orientation density functions (ODFs)

- Lie groups i.e. O(n), SO(n), GL(n), SL(n)
- Kendall shape manifolds

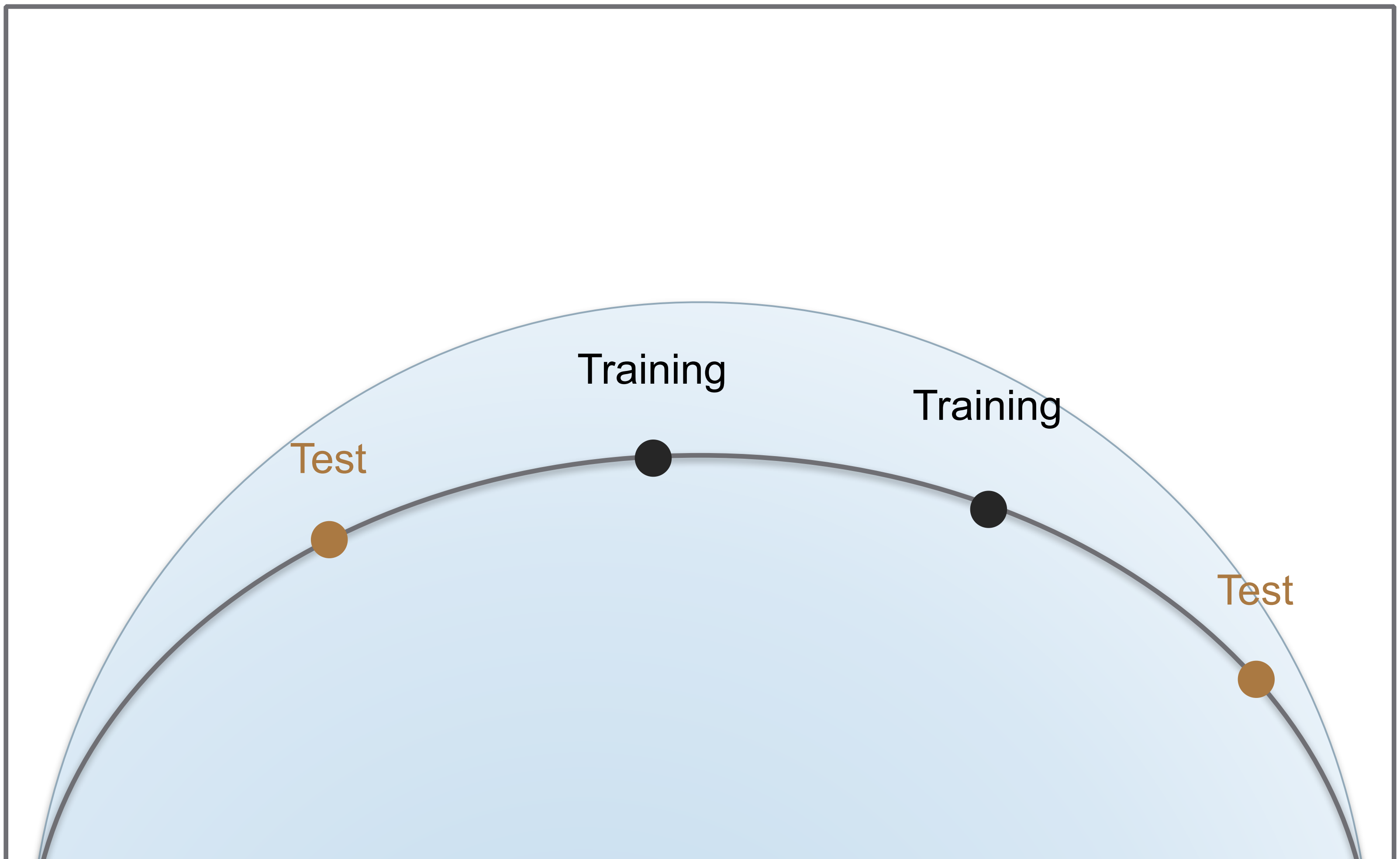
WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

Can we directly use the ordinary general linear model with manifold data?

Euclidean model for manifolds



Euclidean model for manifolds

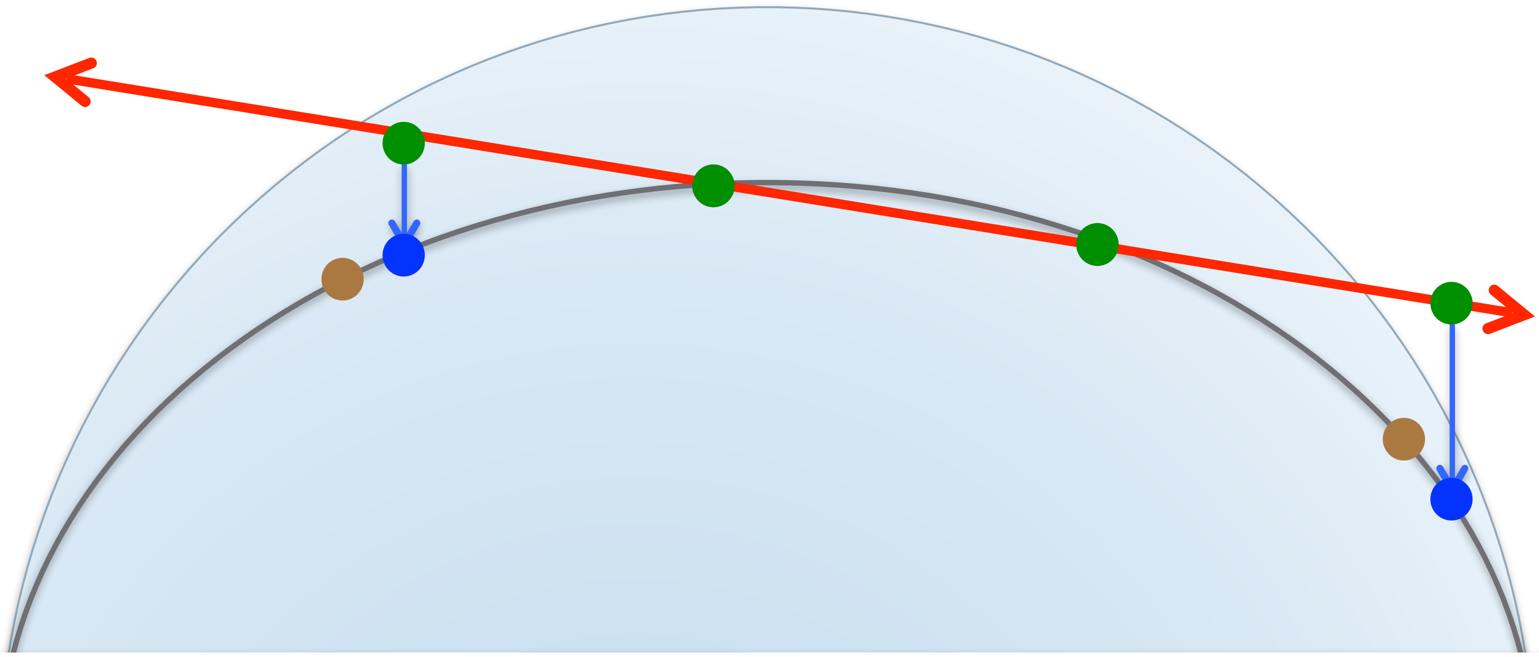


Euclidean model for manifolds

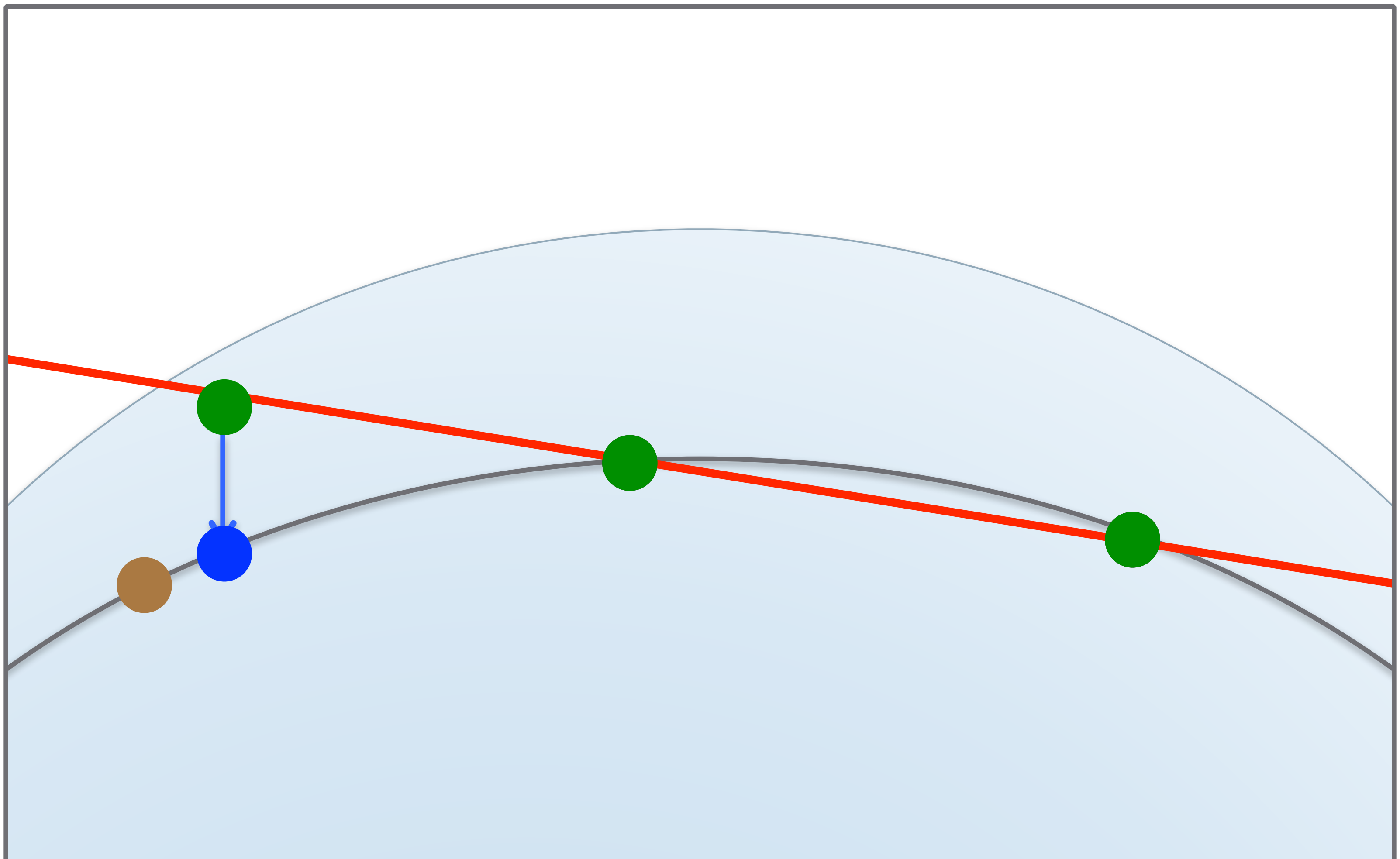
Problem 1: Model in ambient space

Predictions in ambient space

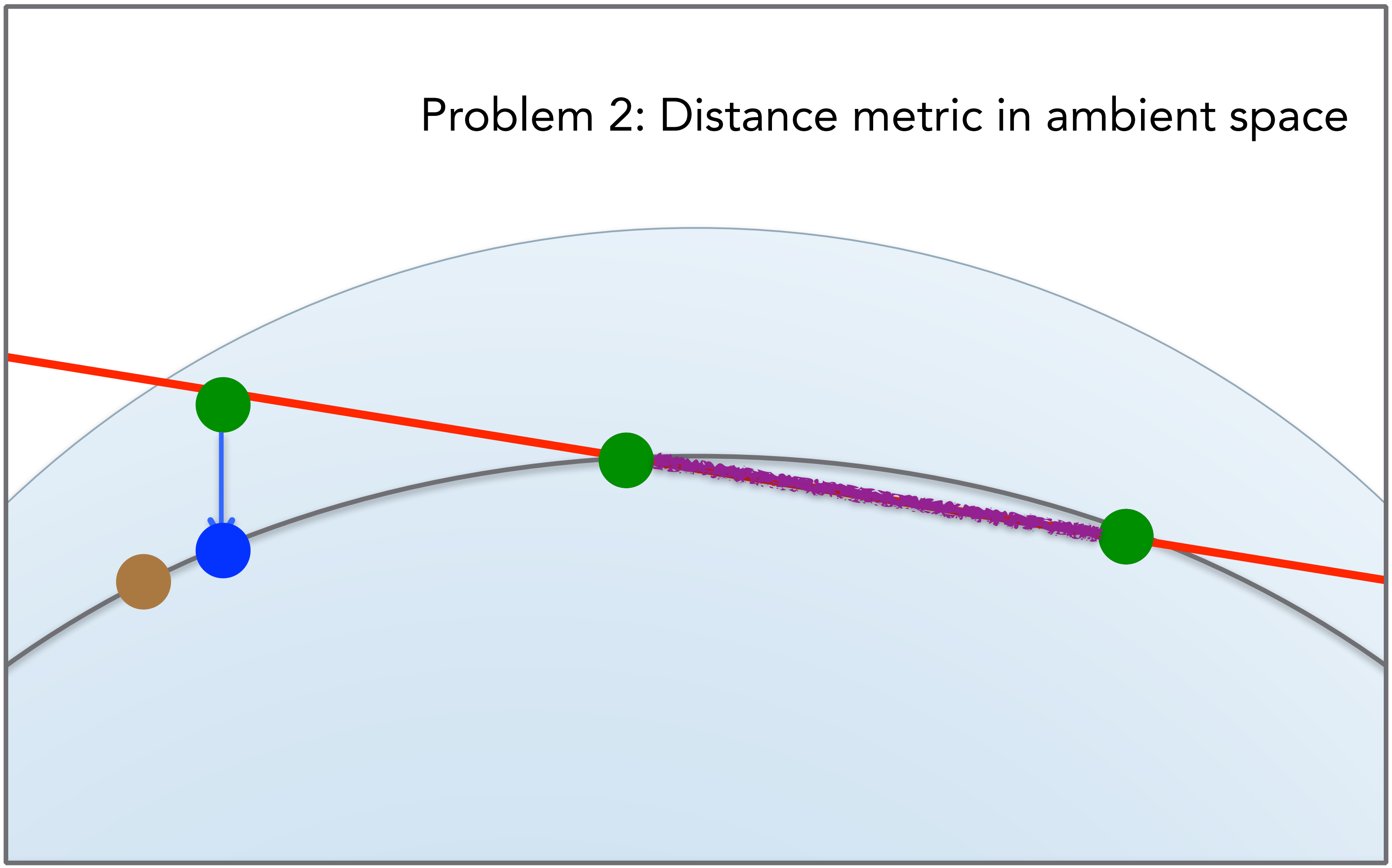
Predictions need to be projected onto manifolds



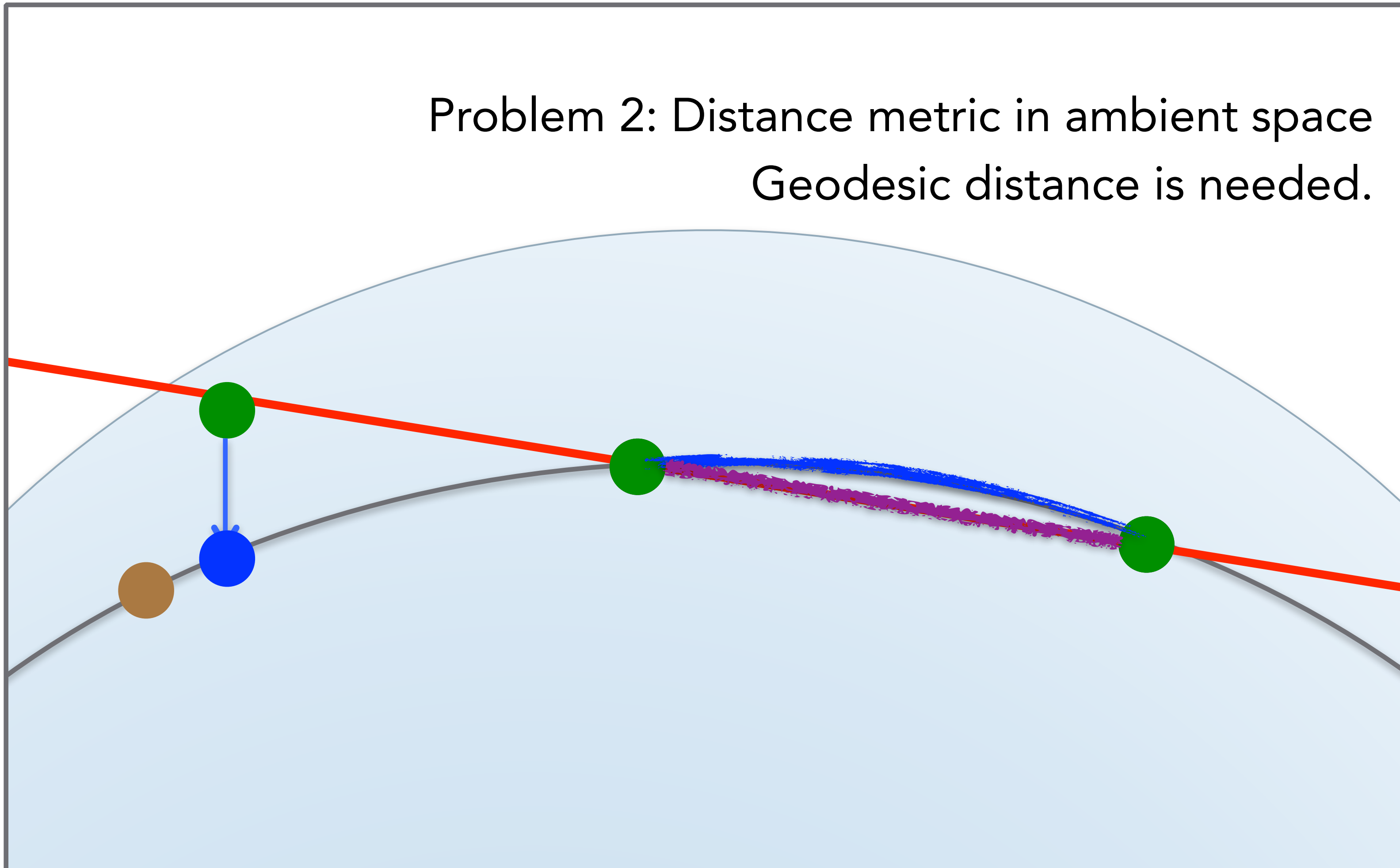
Euclidean model for manifolds



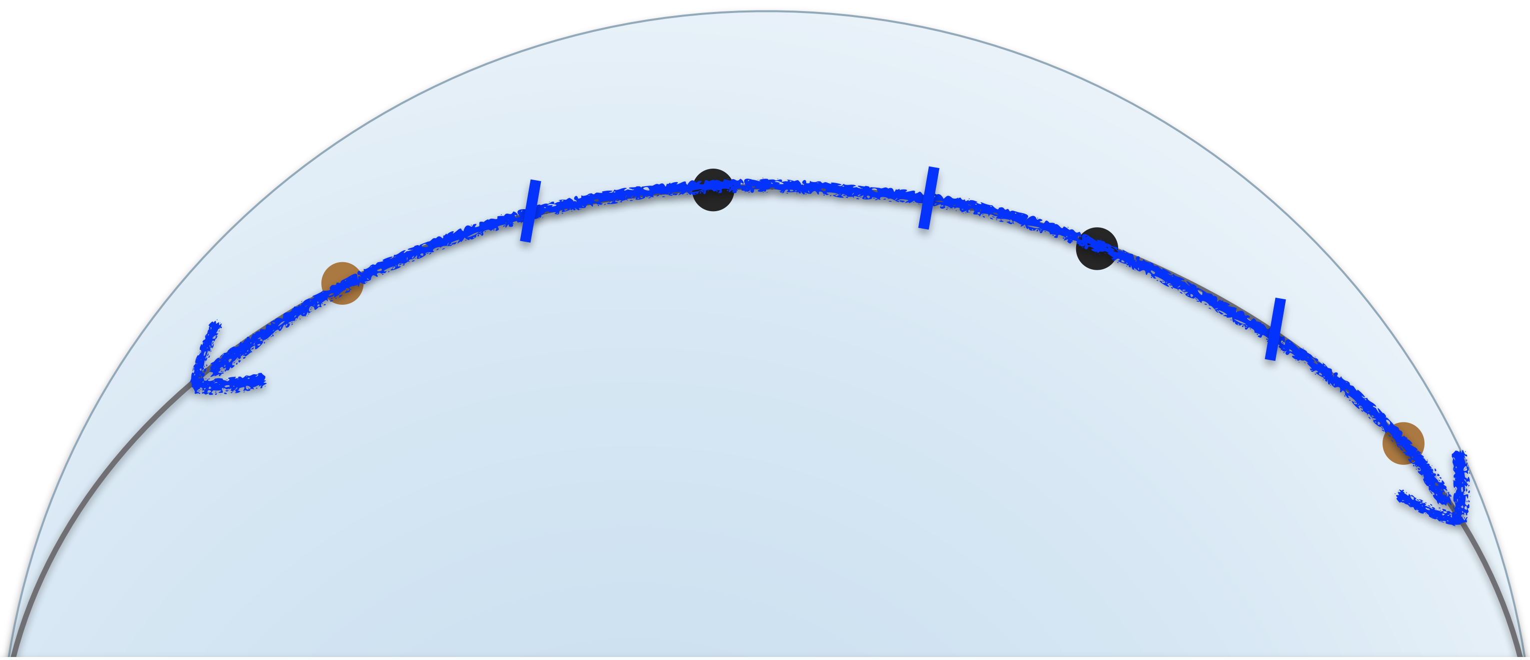
Euclidean model for manifolds



Euclidean model for manifolds



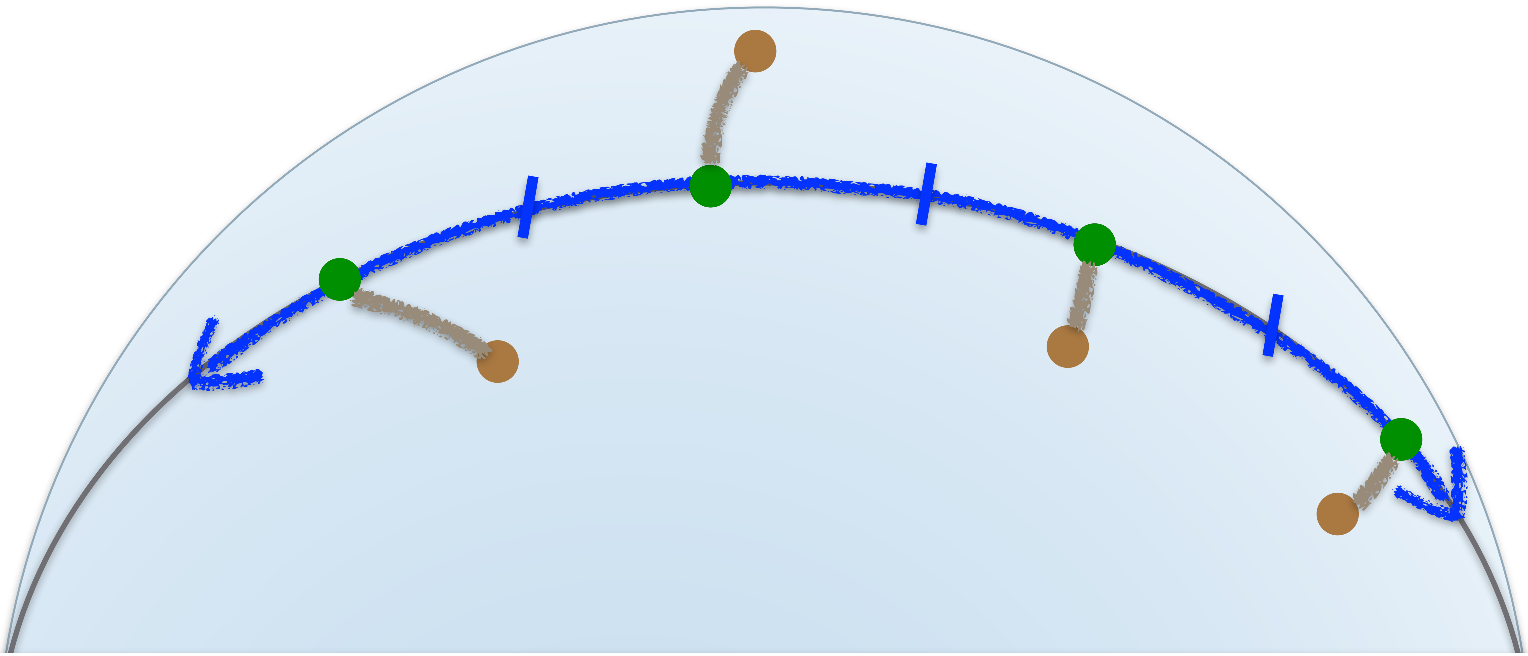
Model on manifolds



Regression with a Single Covariate

$f : R \to \mathcal{M}$

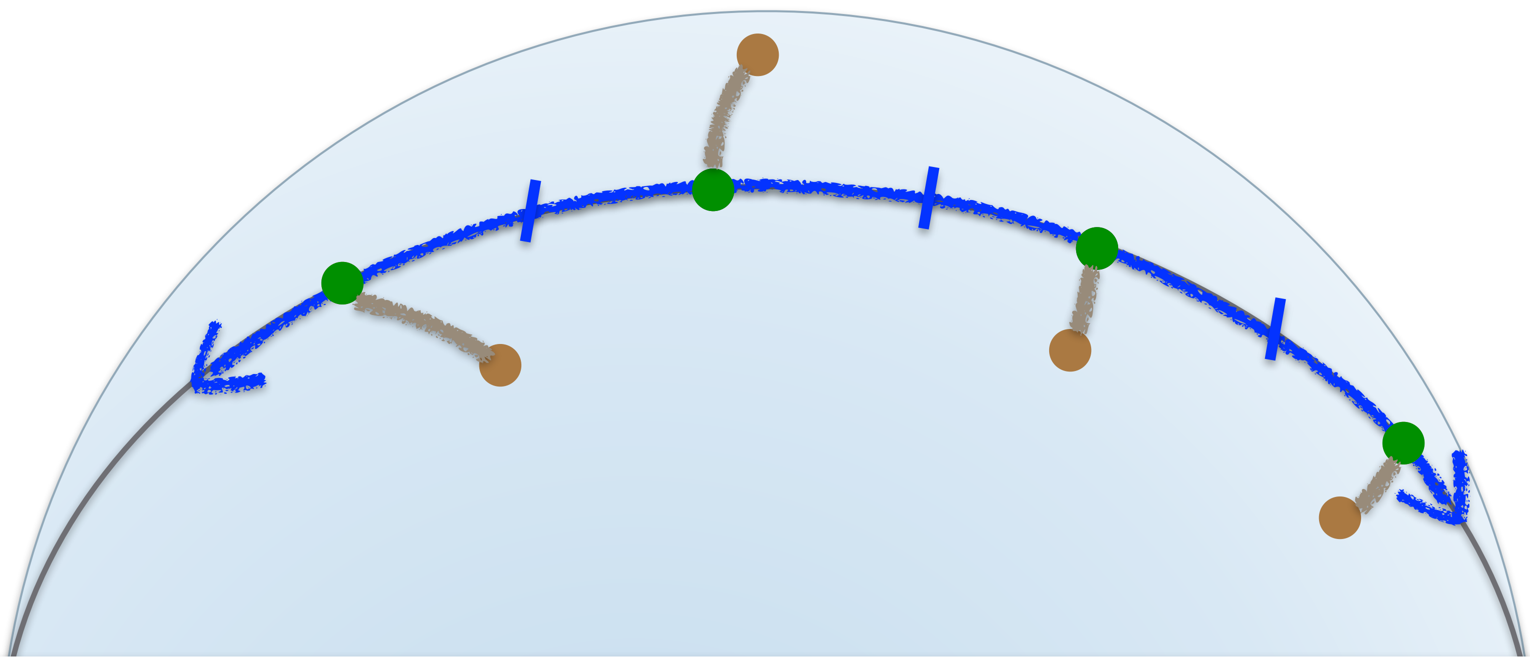
Fletcher, P. Thomas. "Geodesic regression and the theory of least squares on Riemannian manifolds." IJCV, 2013.



Model on manifolds

$f : R \to \mathcal{M}$

Jia Du, Alvina Goh, Sergey Kushnareva, Anqi Qiu, "Geodesic regression on orientation distribution functions with its application to an aging study." NeuroImage, 2014.



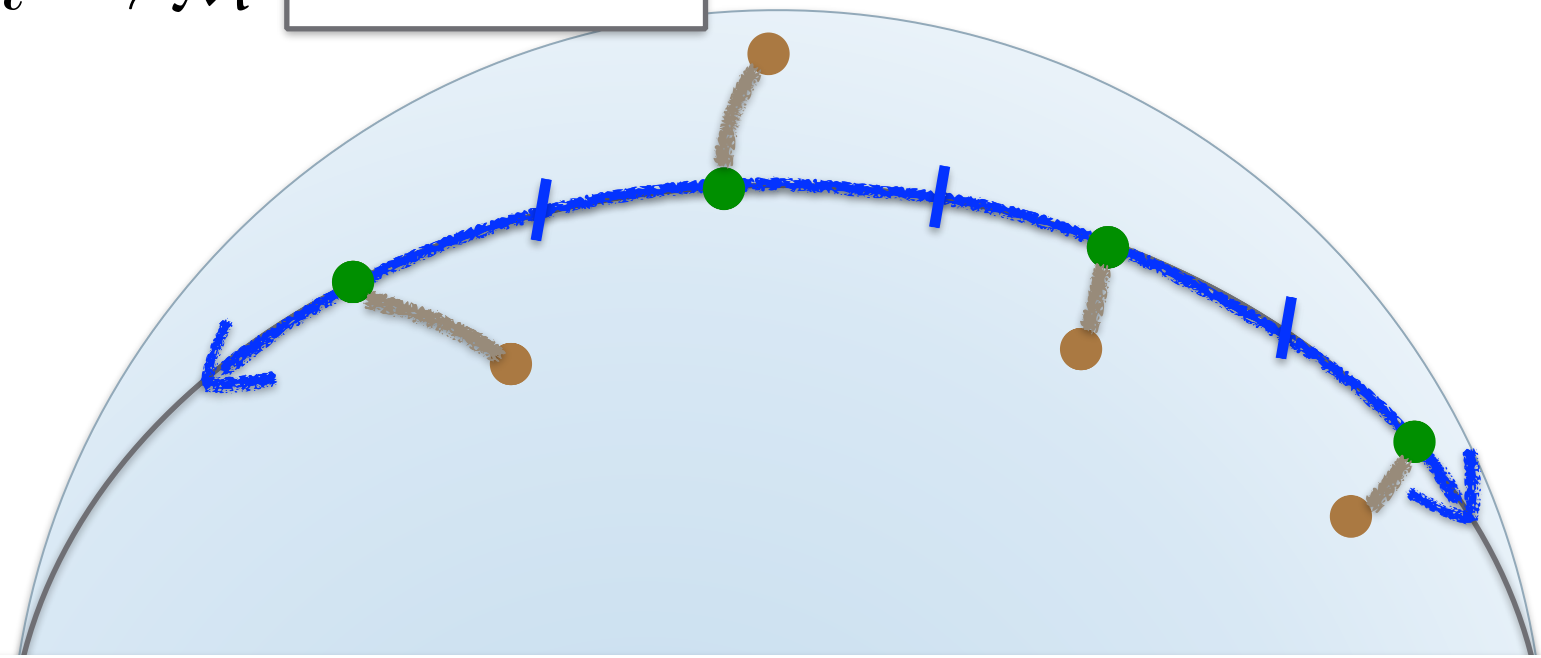
Model on manifolds

$f : R \rightarrow \mathcal{M}$

Jia Du, Alvina Goh, Sergey Kushnareva, Anqi Qiu, "Geodesic regression on orientation distribution functions with its application to an aging study." NeuroImage, 2014.

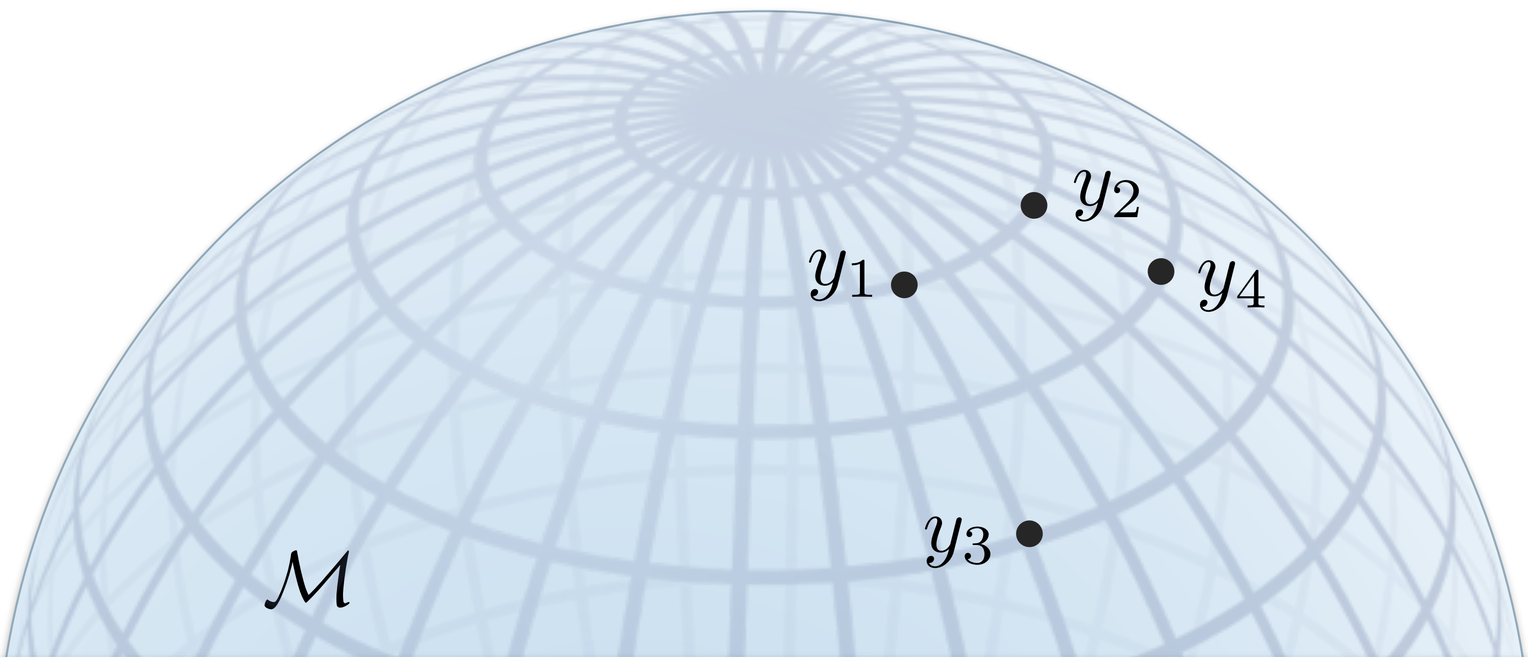
$f : R^n \rightarrow \mathcal{M}$

Our MGLM



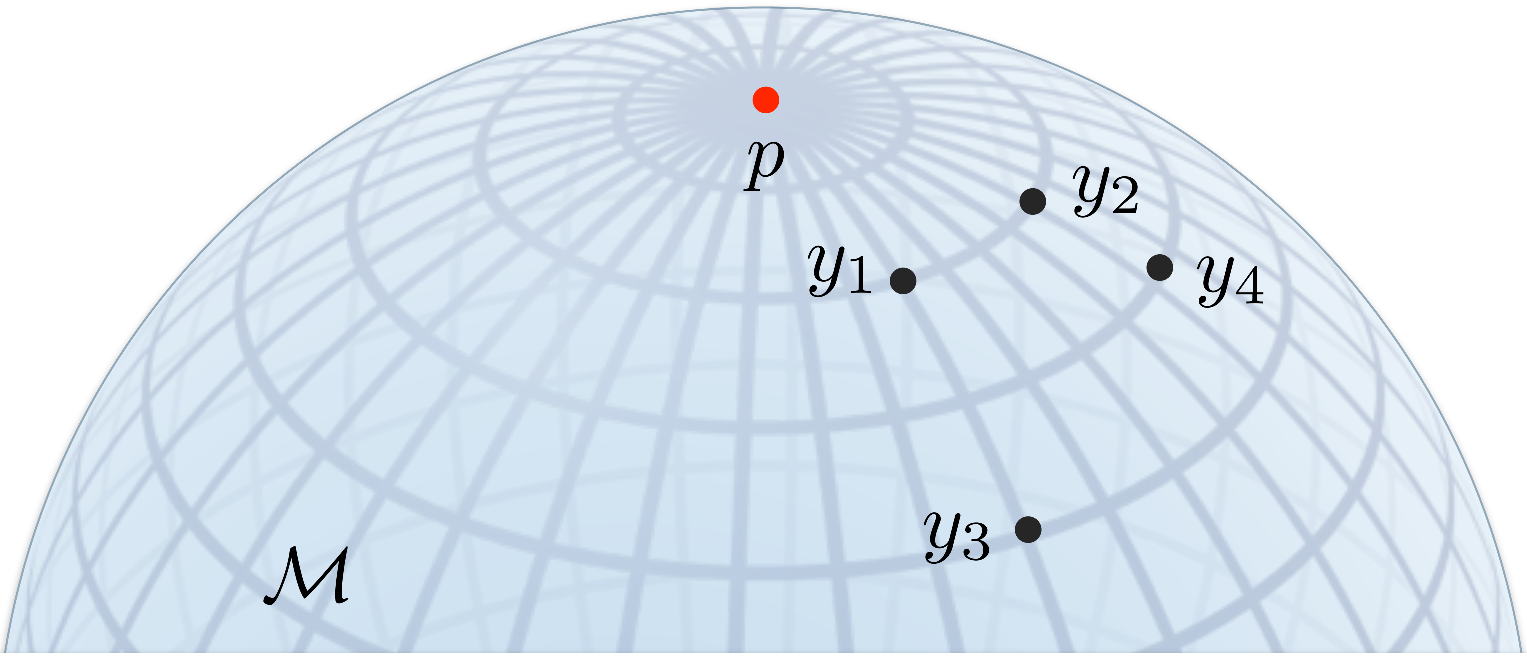
Single covariate

$$\mathrm{E}(p, v) = \frac{1}{2}\sum_{i=1}^{N} \mathrm{d}(\mathrm{Exp}(p, x_i v), y_i)^2$$



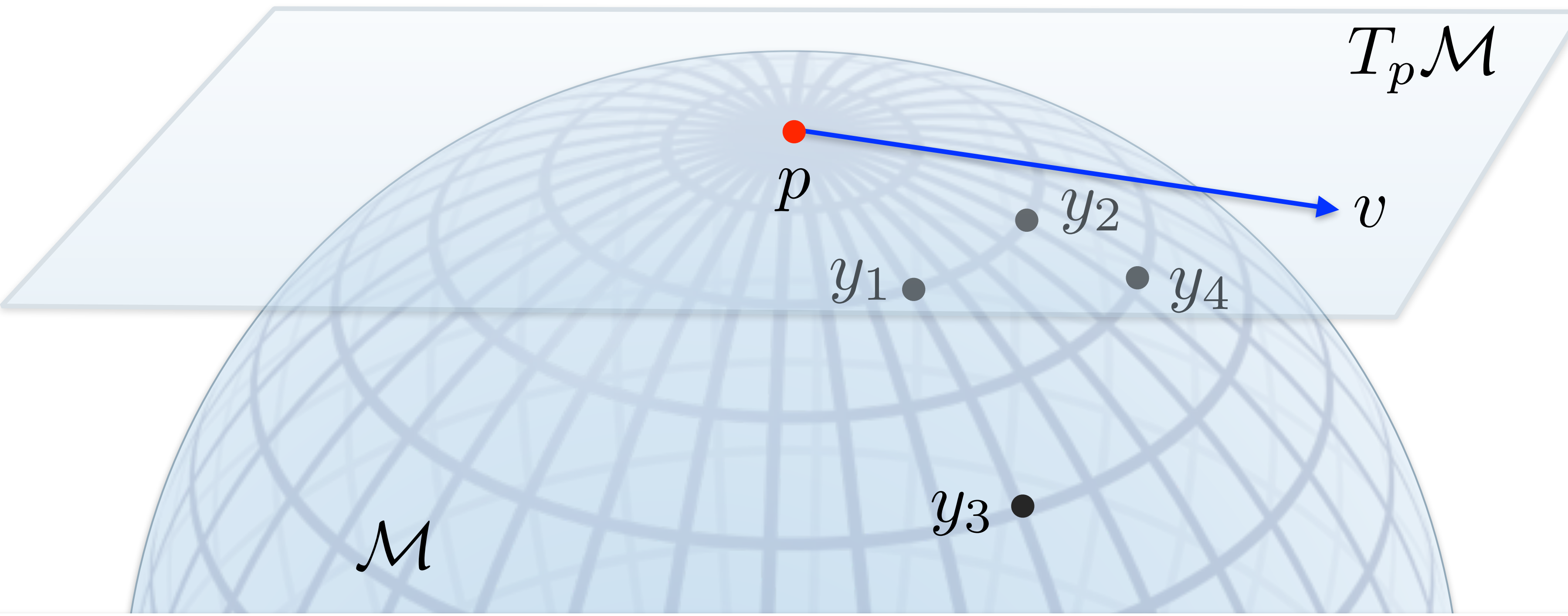
Single covariate

$$\mathrm{E}(p, v) = \frac{1}{2}\sum_{i=1}^{N} \mathrm{d}(\mathrm{Exp}(\underline{p}, x_i v), y_i)^2$$



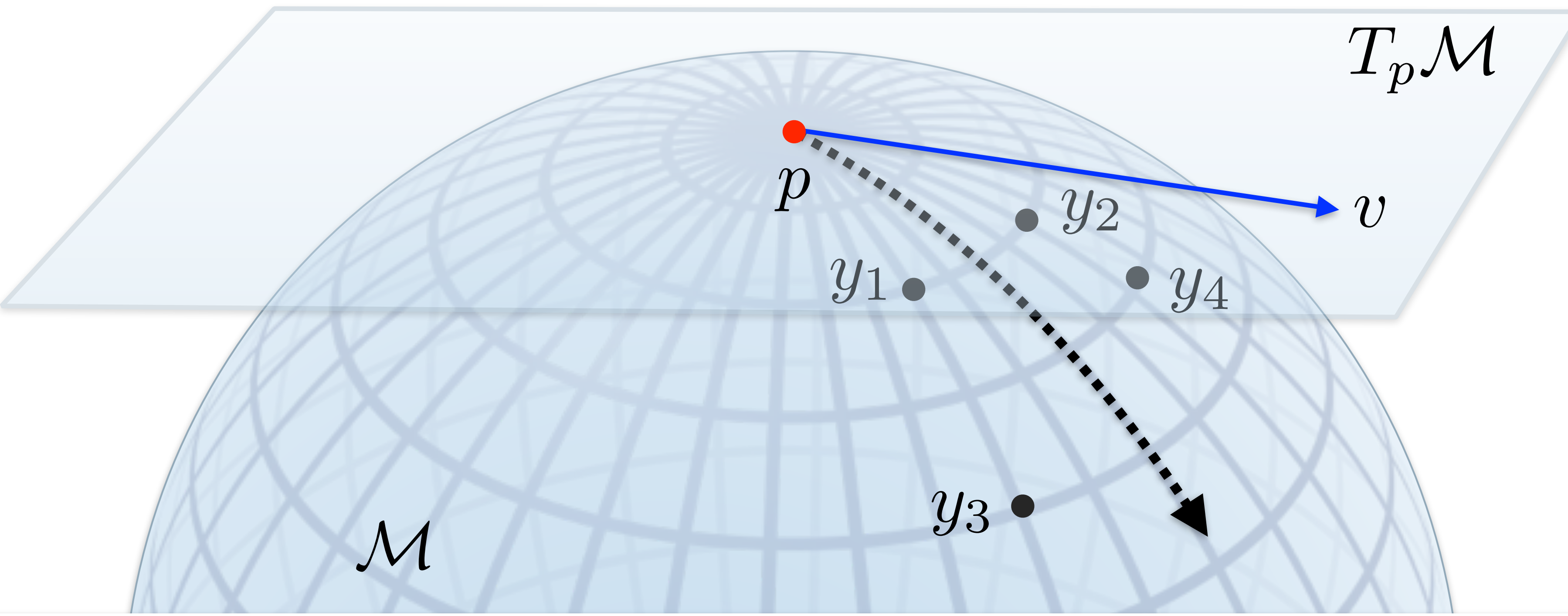
Single covariate

$$\mathrm{E}(p, v) = \frac{1}{2} \sum_{i=1}^{N} \mathrm{d}(\mathrm{Exp}(p, x_i \underline{v}), y_i)^2$$



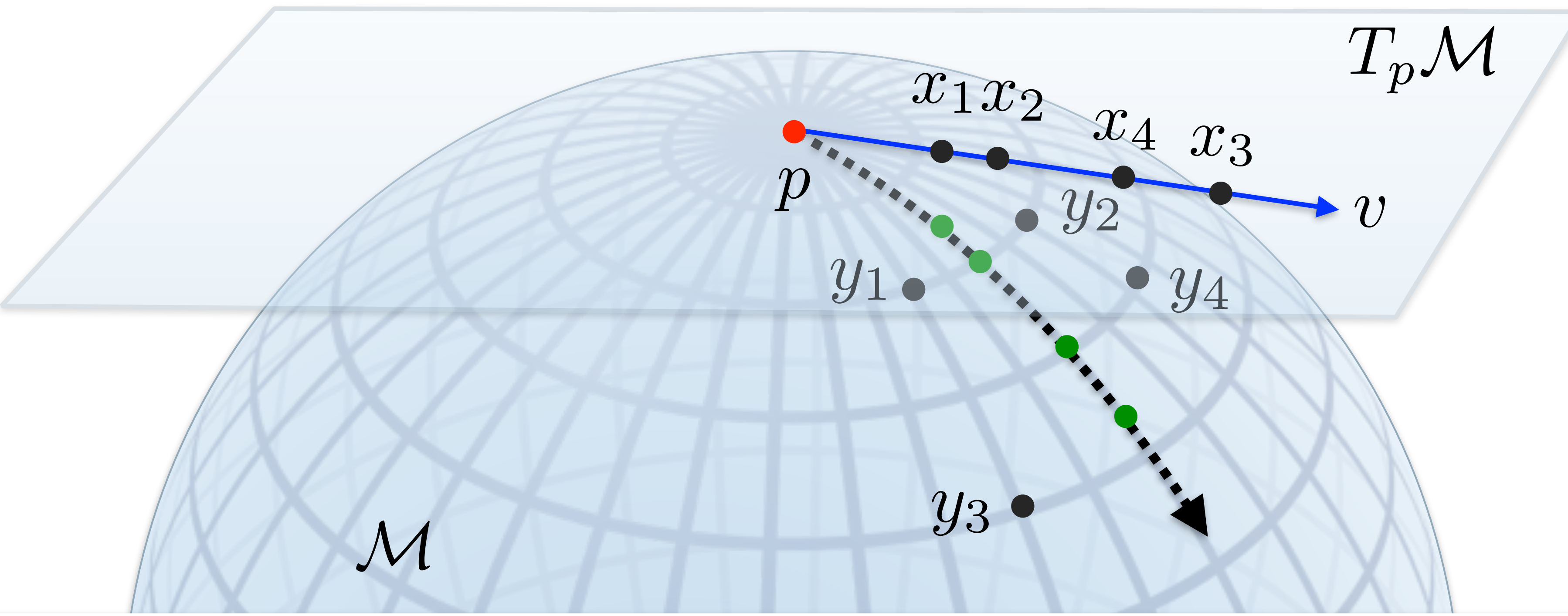
Single covariate

$$\mathrm{E}(p, v) = \frac{1}{2} \sum_{i=1}^{N} \mathrm{d}(\underline{\mathrm{Exp}(p, x_i v)}, y_i)^2$$



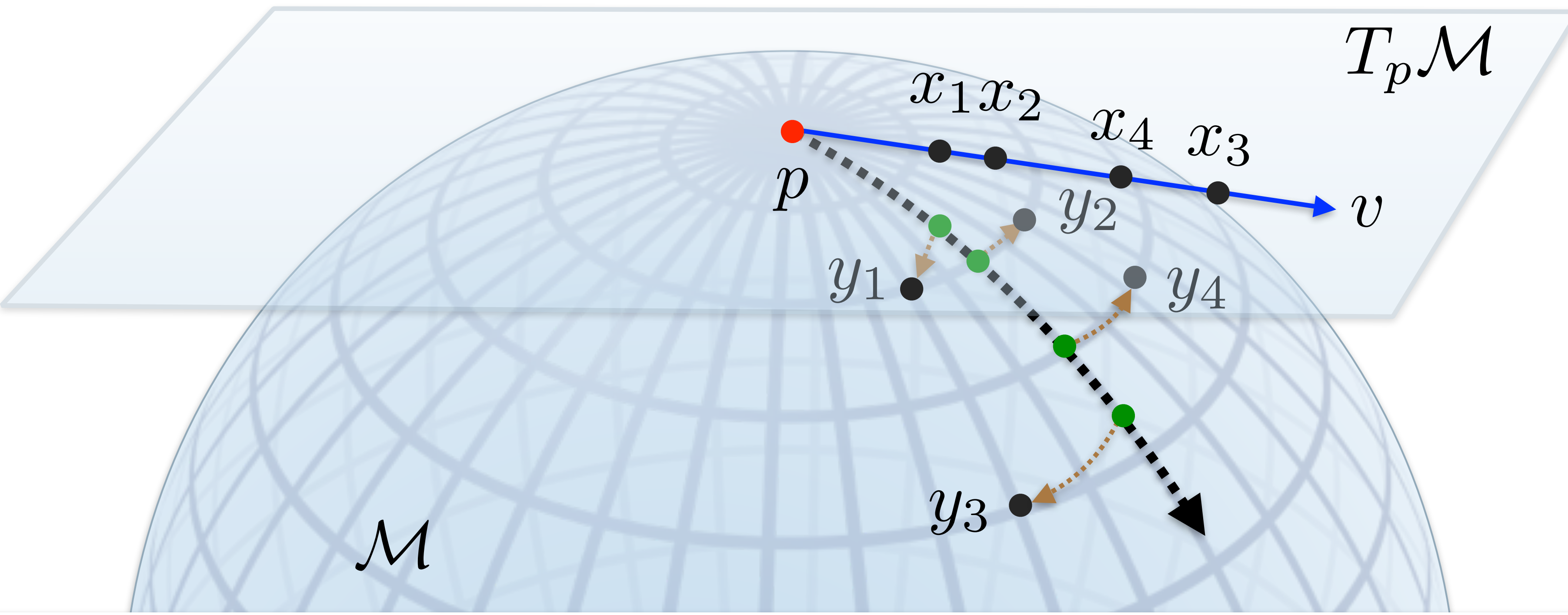
Single covariate

$$\mathrm{E}(p, v) = \frac{1}{2} \sum_{i=1}^{N} \mathrm{d}(\underline{\mathrm{Exp}(p, x_i v)}, y_i)^2$$



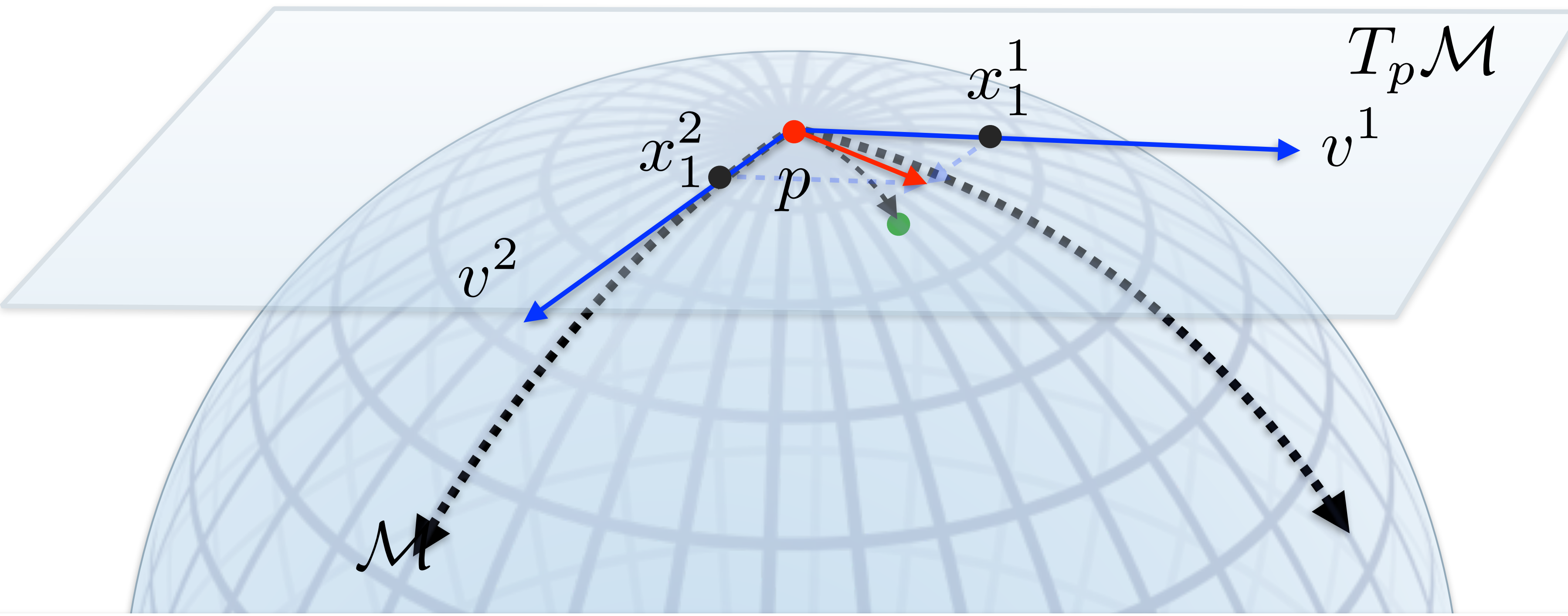
Single covariate

$$\mathrm{E}(p, v) = \frac{1}{2} \sum_{i=1}^{N} \underline{\mathrm{d}(\mathrm{Exp}(p, x_i v), y_i)^2}$$



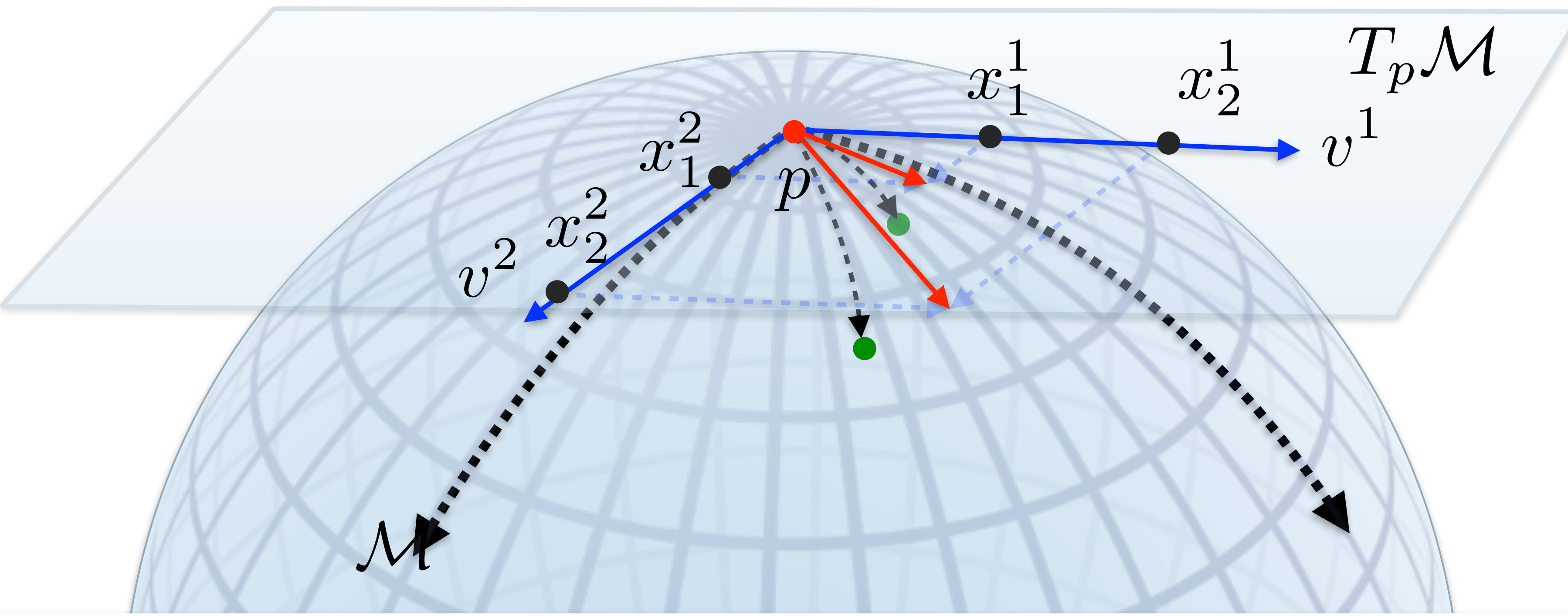
Multiple covariates

$$\mathrm{E}(p, \boldsymbol{V}) = \frac{1}{2}\sum_{i=1}^{N} \mathrm{d}(\mathrm{Exp}(p, \sum_j x_i^j v^j), y_i)^2$$



Multiple covariates

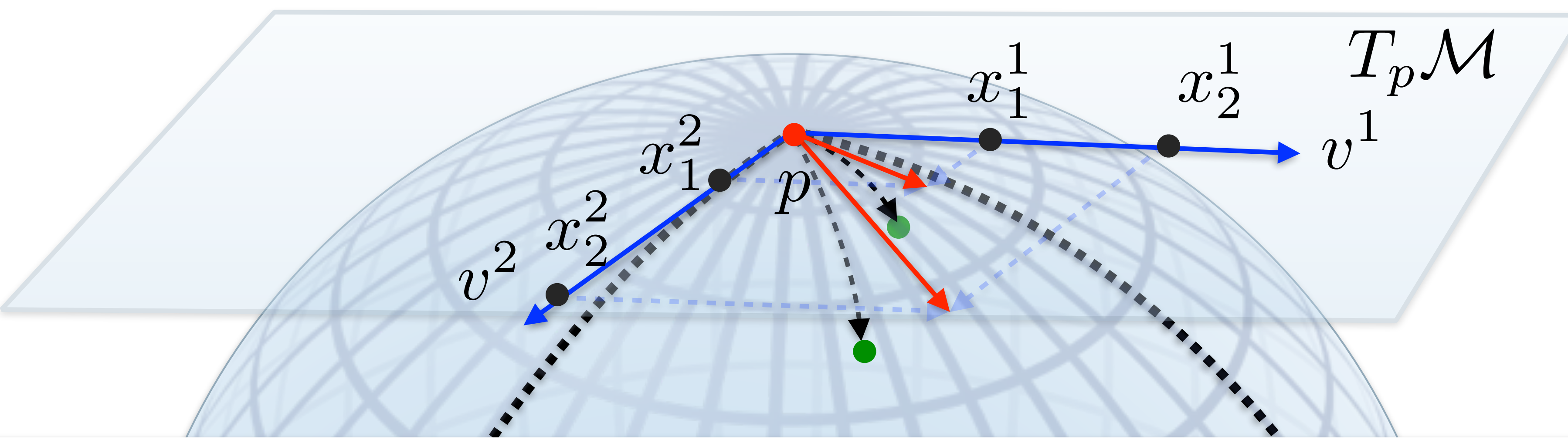
$$\mathrm{E}(p, \boldsymbol{V}) = \frac{1}{2}\sum_{i=1}^{N} \mathrm{d}(\mathrm{Exp}(p, \sum_{j} x_i^j v^j), y_i)^2$$



Multiple covariates

$$E(p, \boldsymbol{V}) = \frac{1}{2}\sum_{i=1}^{N} d(\mathrm{Exp}(p, \sum_j x_i^j v^j), y_i)^2$$

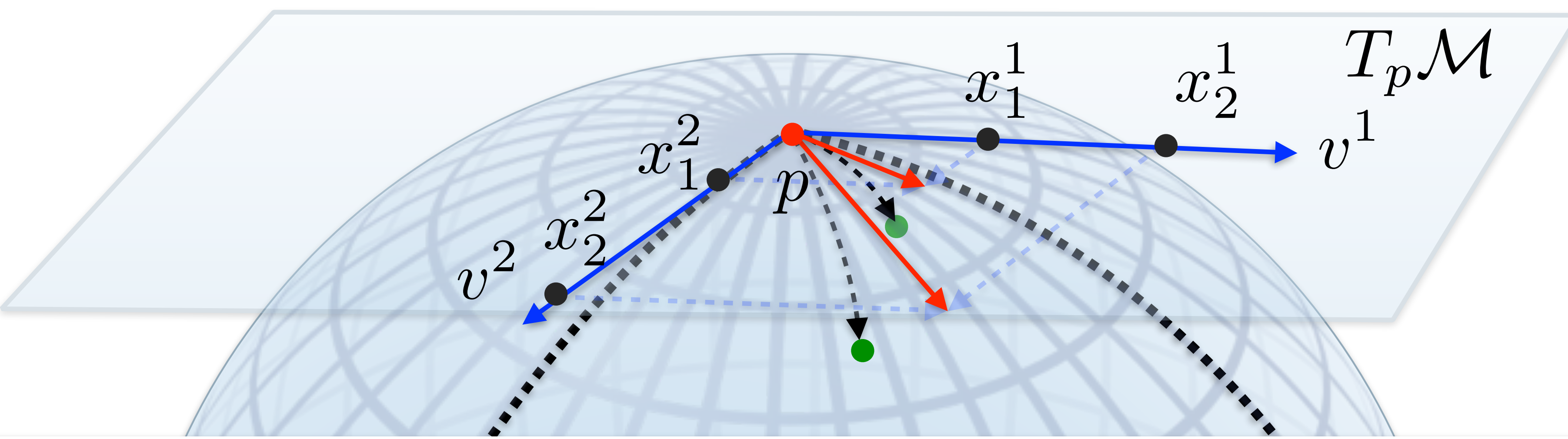
$$y = \alpha + \beta^1 x^1 + \beta^2 x^2 + \beta^3 x^3 + \epsilon$$



Multiple covariates

$$\mathrm{E}(p, \boldsymbol{V}) = \frac{1}{2}\sum_{i=1}^{N} \mathrm{d}(\mathrm{Exp}(\underline{p}, \sum_j x_i^j v^j), y_i)^2$$

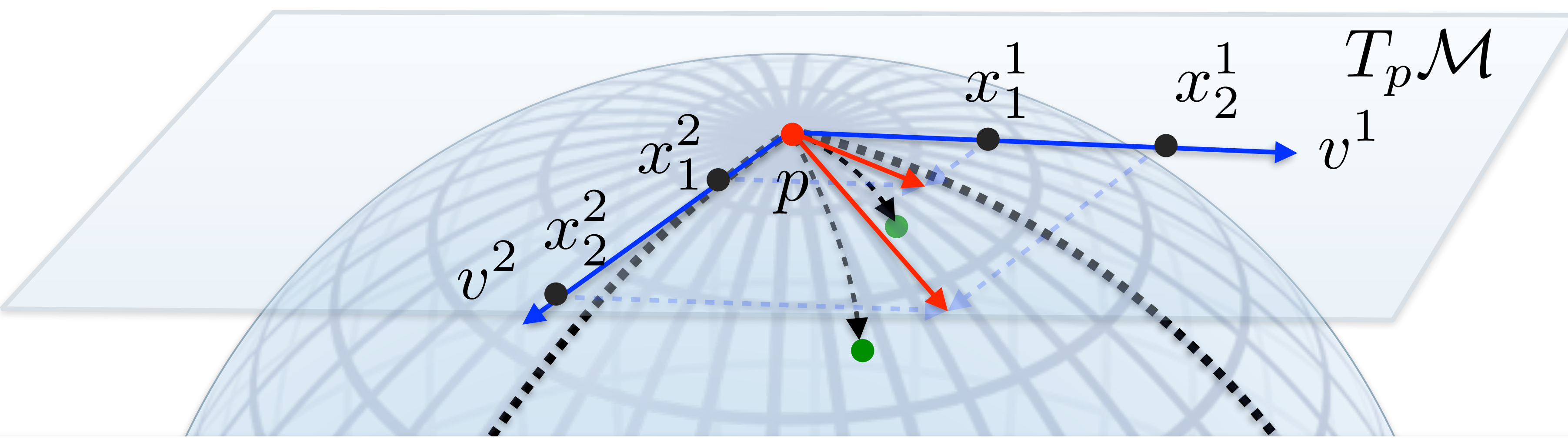
$$y = \underline{\alpha} + \beta^1 x^1 + \beta^2 x^2 + \beta^3 x^3 + \epsilon$$



Multiple covariates

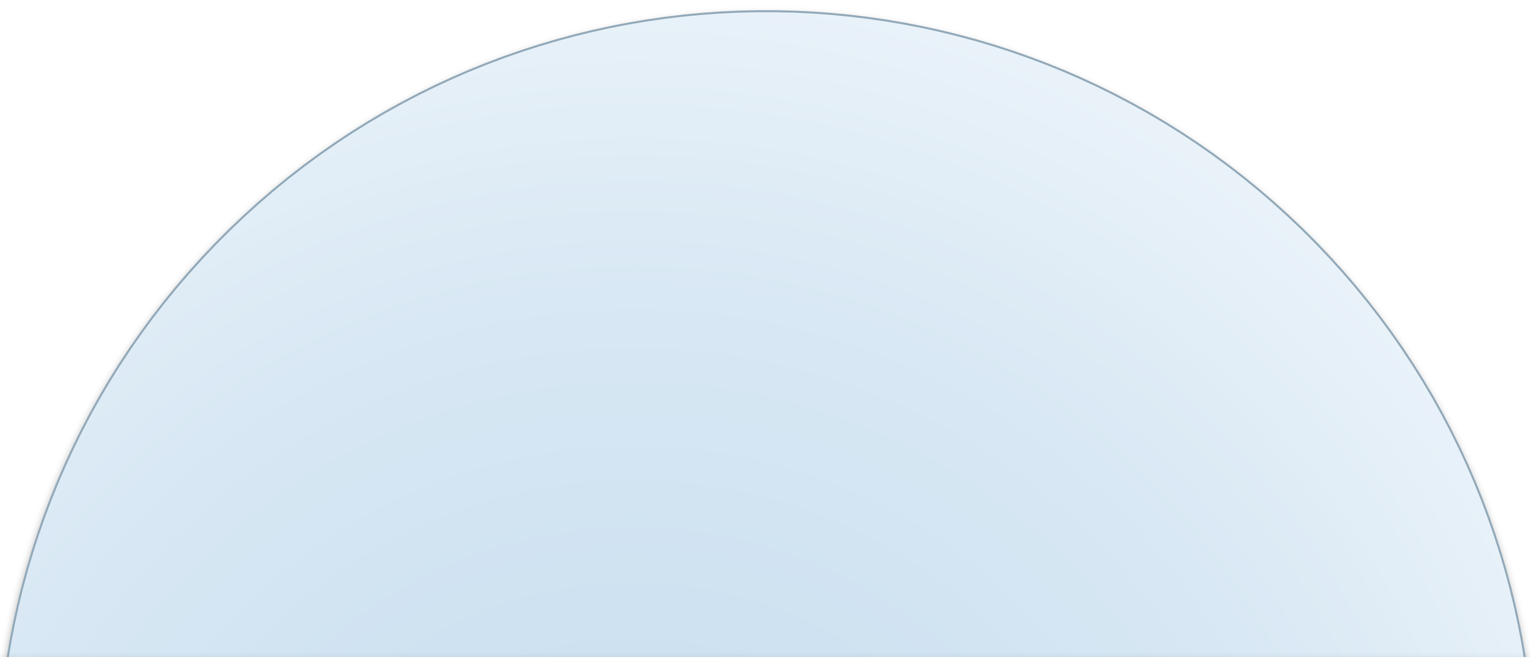
$$\mathrm{E}(p, \boldsymbol{V}) = \frac{1}{2} \sum_{i=1}^{N} \mathrm{d}(\mathrm{Exp}(\underline{p}, \sum_j x_i^j \underline{v^j}), y_i)^2$$

$$y = \underline{\alpha} + \underline{\beta^1} x^1 + \underline{\beta^2} x^2 + \underline{\beta^3} x^3 + \epsilon$$



Optimization (iterative method)

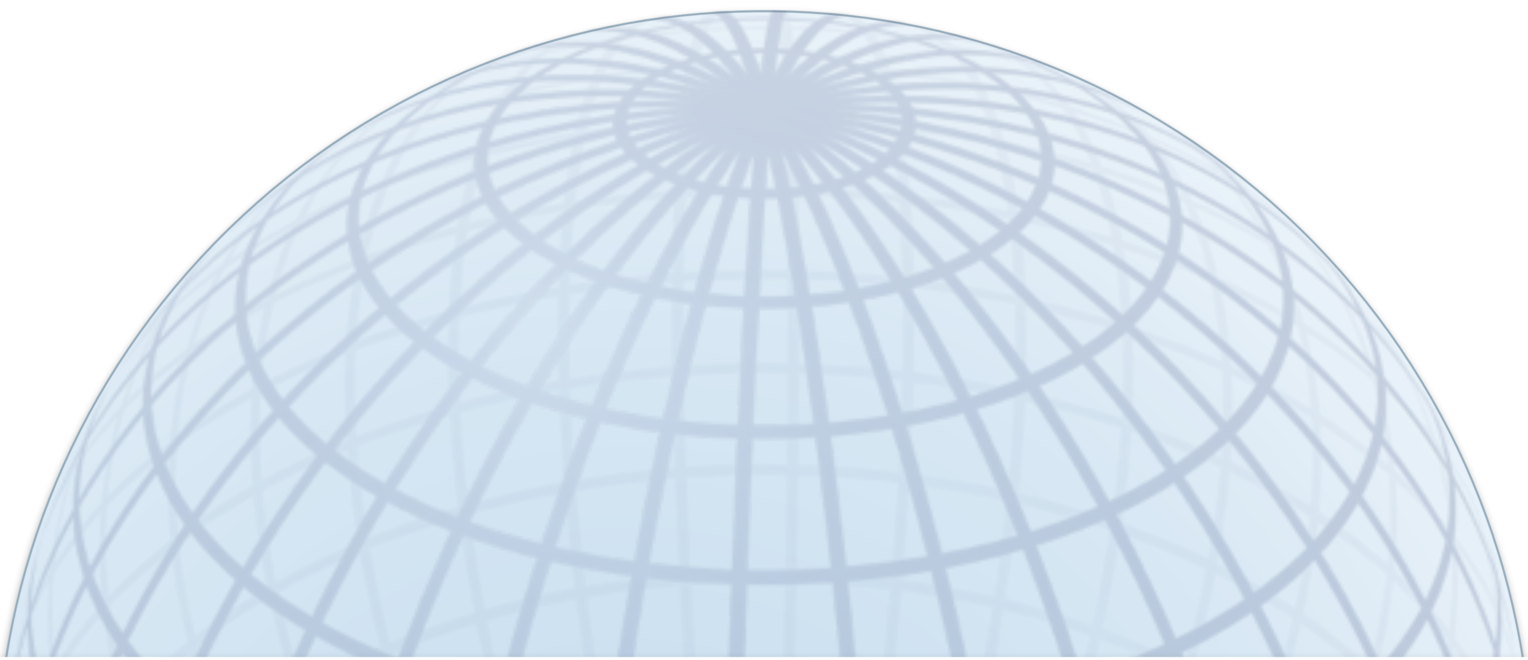
$$\mathrm{E}(p, \boldsymbol{V}) = \frac{1}{2} \sum_{i=1}^{N} \mathrm{d}(\mathrm{Exp}(p, \sum_{j} x_i^j v^j), y_i)^2$$



Optimization (iterative method)

$$\mathrm{E}(p, \boldsymbol{V}) = \frac{1}{2} \sum_{i=1}^{N} \mathrm{d}(\mathrm{Exp}(p, \sum_j x_i^j v^j), y_i)^2$$

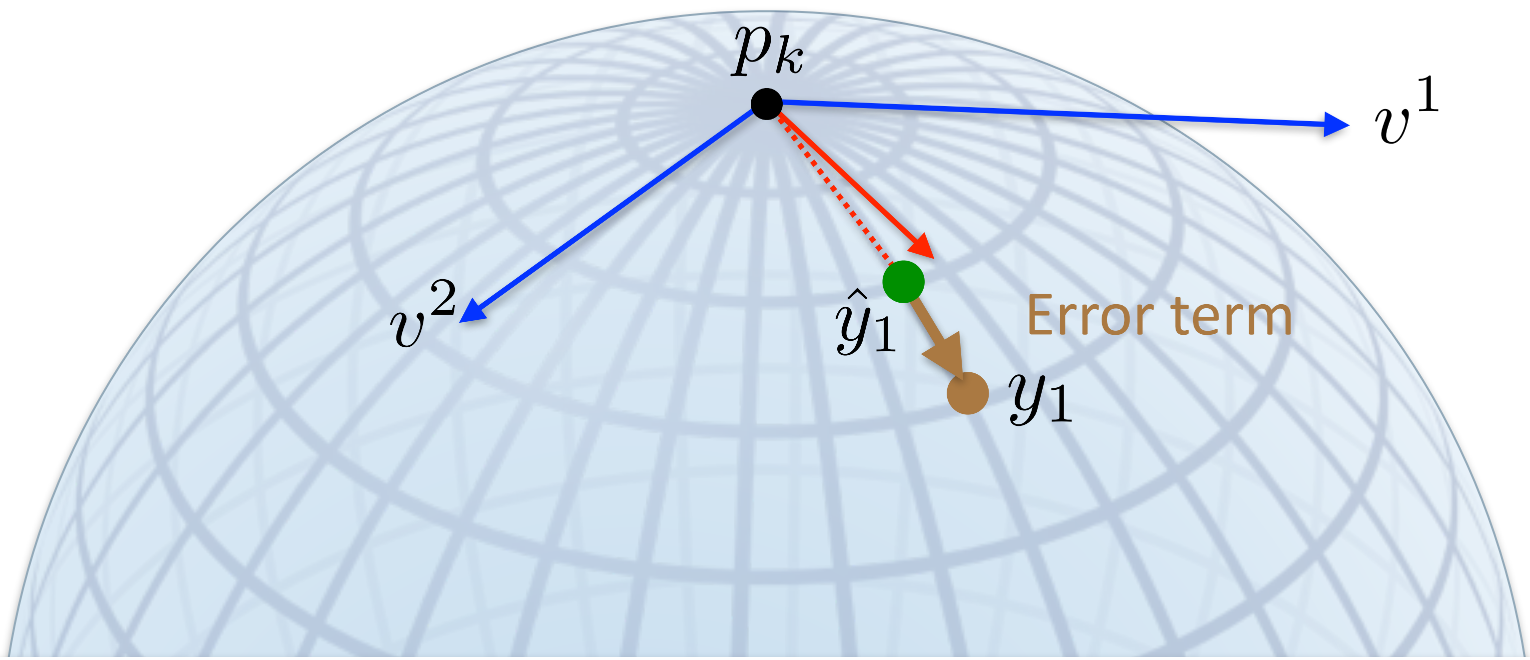
Step 1 : $p_{k+1} \leftarrow \mathrm{Exp}\left(p_k, \alpha \Sigma_i \Gamma_{\hat{y}_i \to p_k} \mathrm{Log}(\hat{y}_i, y_i)\right)$



Optimization (iterative method)

$$\mathrm{E}(p, \boldsymbol{V}) = \frac{1}{2}\sum_{i=1}^{N} \mathrm{d}(\mathrm{Exp}(p, \sum_j x_i^j v^j), y_i)^2$$

Step 1 : $p_{k+1} \leftarrow \mathrm{Exp}\left(p_k, \alpha \Sigma_i \Gamma_{\hat{y}_i \to p_k} \underline{\mathrm{Log}(\hat{y}_i, y_i)}\right)$ Error term

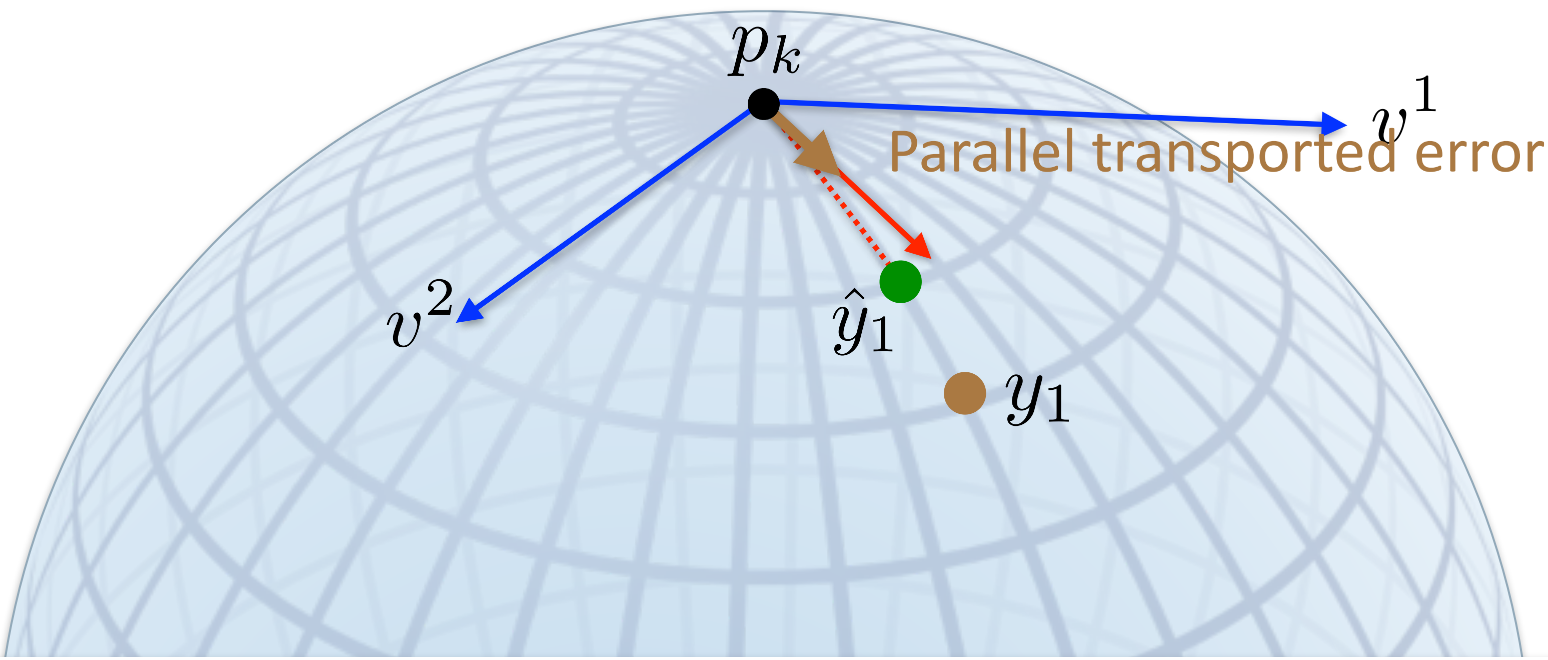


Optimization (iterative method)

$$\mathrm{E}(p, \boldsymbol{V}) = \frac{1}{2} \sum_{i=1}^{N} \mathrm{d}(\mathrm{Exp}(p, \sum_j x_i^j v^j), y_i)^2$$

Step 1 : $p_{k+1} \leftarrow \mathrm{Exp}\left(p_k, \alpha \Sigma_i \underline{\Gamma_{\hat{y}_i \to p_k} \mathrm{Log}(\hat{y}_i, y_i)}\right)$

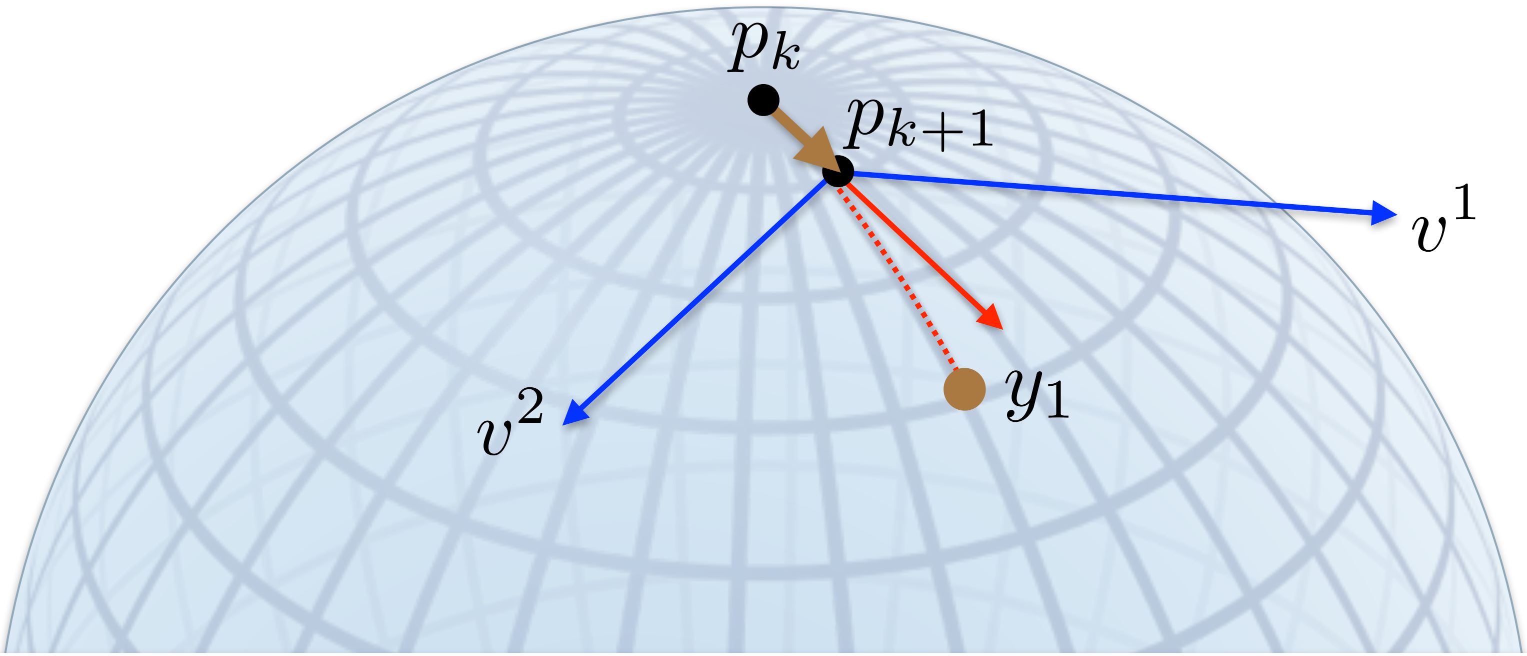
Parallel transported error



Optimization (iterative method)

$$\mathrm{E}(p, \boldsymbol{V}) = \frac{1}{2} \sum_{i=1}^{N} \mathrm{d}(\mathrm{Exp}(p, \sum_j x_i^j v^j), y_i)^2$$

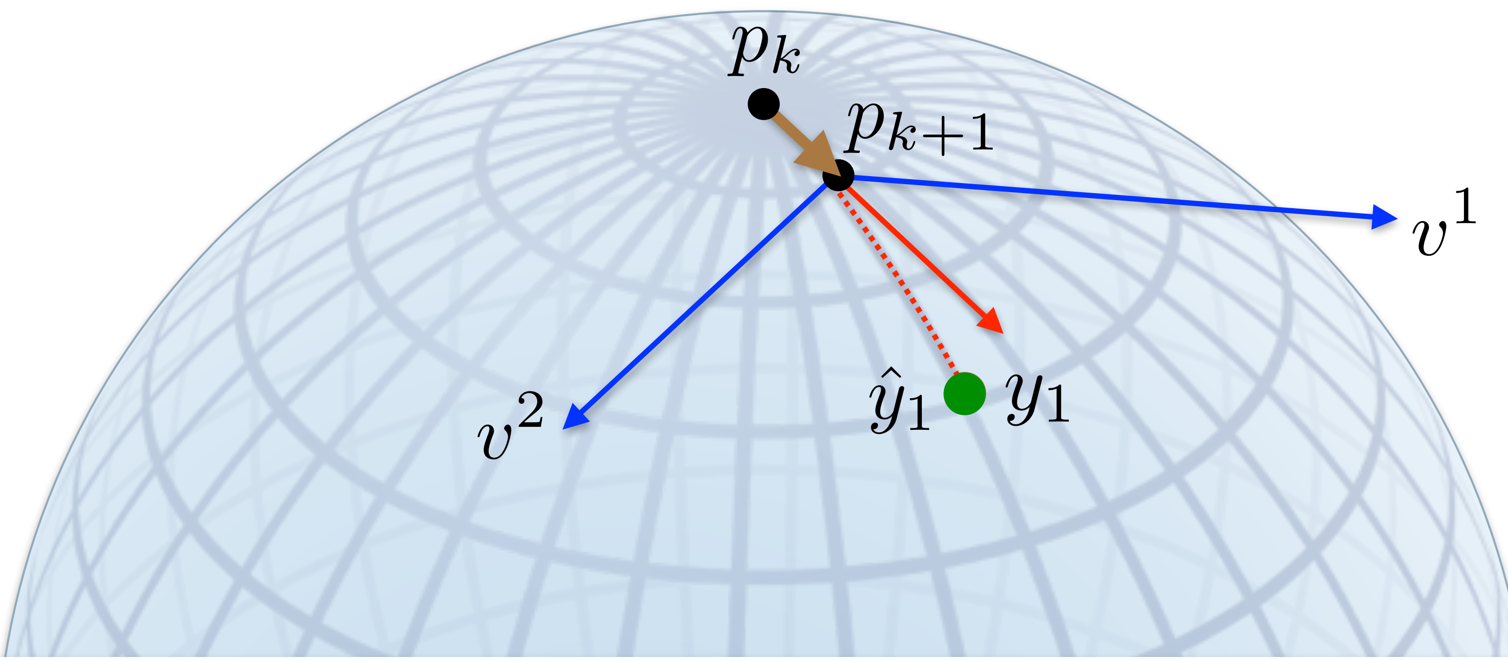
Step 1 : $p_{k+1} \leftarrow \mathrm{Exp}\left(p_k, \underline{\alpha \Sigma_i \Gamma_{\hat{y}_i \to p_k} \mathrm{Log}(\hat{y}_i, y_i)}\right)$



Optimization (iterative method)

$$\mathrm{E}(p, \boldsymbol{V}) = \frac{1}{2} \sum_{i=1}^{N} \mathrm{d}(\mathrm{Exp}(p, \sum_j x_i^j v^j), y_i)^2$$

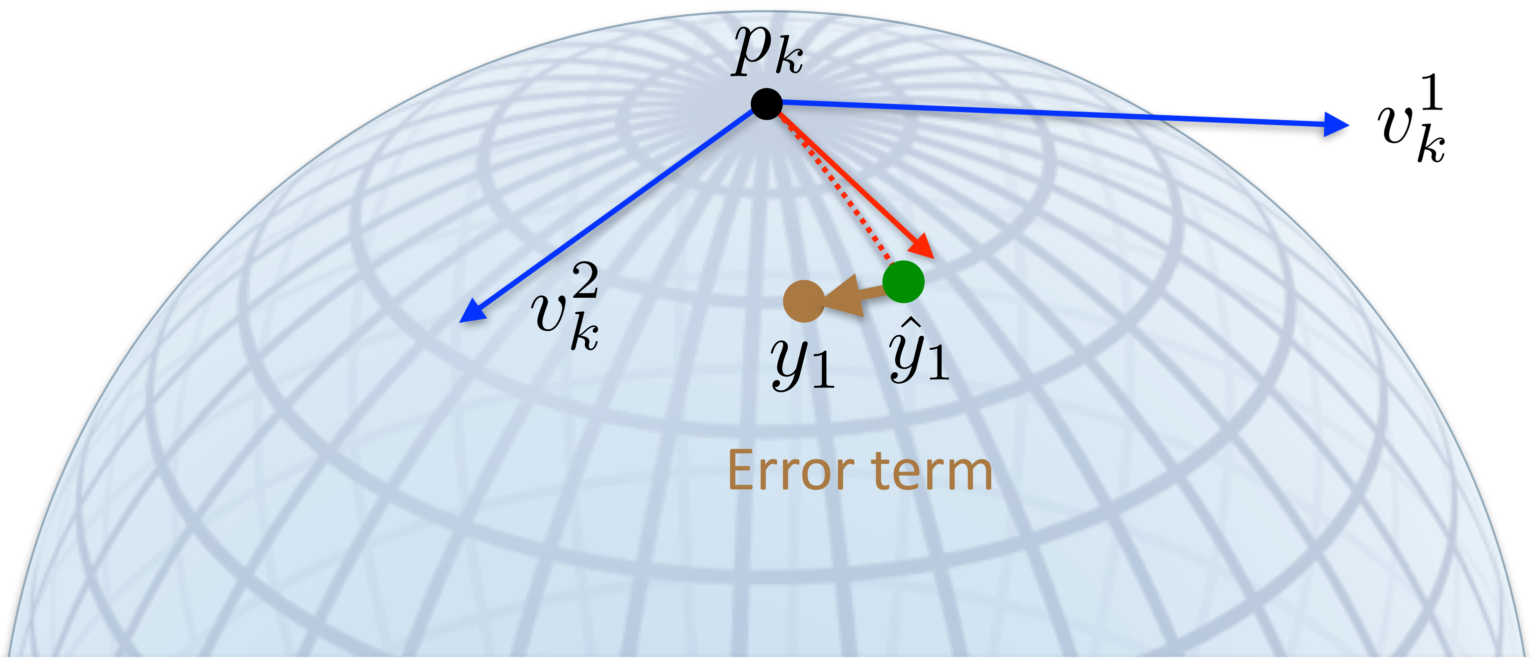
Step 1 : $p_{k+1} \leftarrow \mathrm{Exp}\left(p_k, \underline{\alpha \Sigma_i \Gamma_{\hat{y}_i \to p_k} \mathrm{Log}(\hat{y}_i, y_i)}\right)$



Optimization (iterative method)

$$\mathrm{E}(p, \boldsymbol{V}) = \frac{1}{2} \sum_{i=1}^{N} \mathrm{d}(\mathrm{Exp}(p, \sum_j x_i^j v^j), y_i)^2$$

Step 2 : $v^j_{k+1/2} = v^j_k + \alpha \sum_i x^j_i \Gamma_{\hat{y}_i \to p} \underline{\mathrm{Log}(\hat{y}_i, y_i)}$ Error term

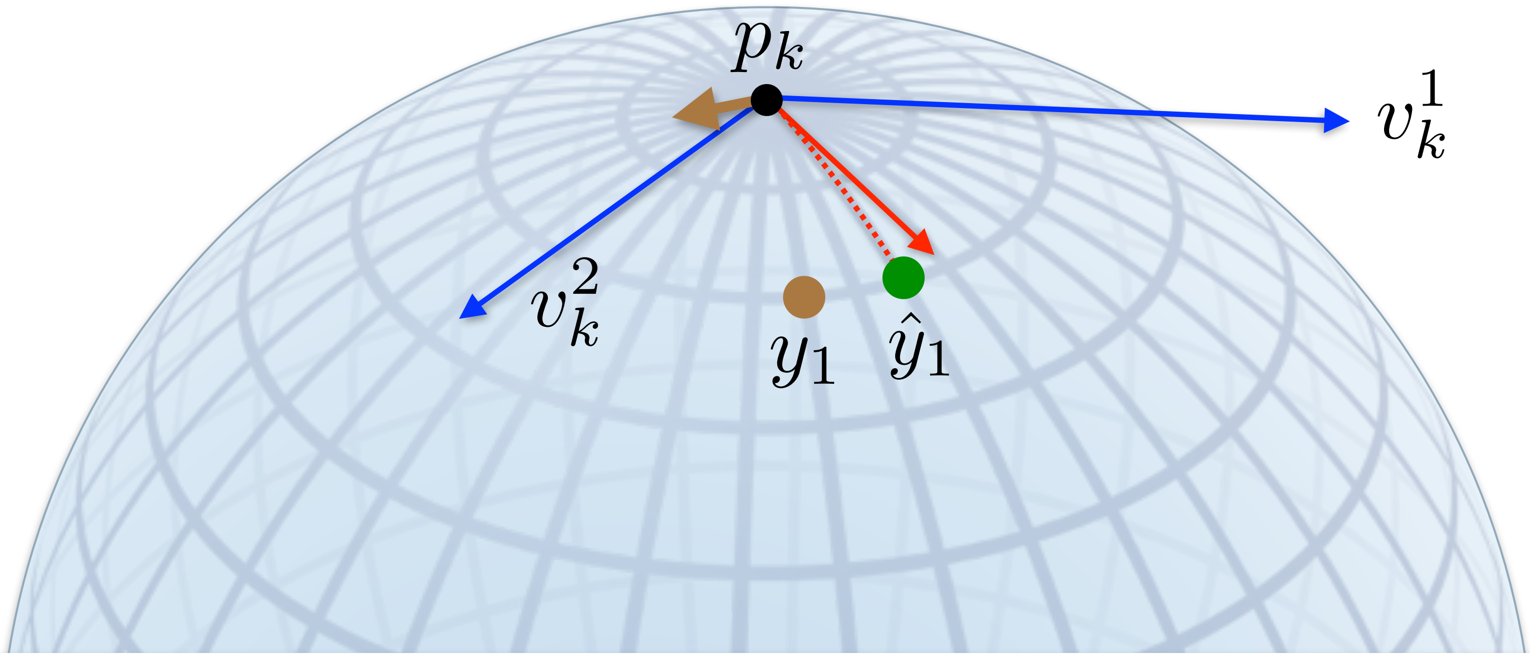


Optimization (iterative method)

$$\mathrm{E}(p, \boldsymbol{V}) = \frac{1}{2} \sum_{i=1}^{N} \mathrm{d}(\mathrm{Exp}(p, \sum_j x_i^j v^j), y_i)^2$$

Step 2 : $$v_{k+1/2}^j = v_k^j + \alpha \sum_i x_i^j \underline{\Gamma_{\hat{y}_i \to p} \mathrm{Log}(\hat{y}_i, y_i)}$$

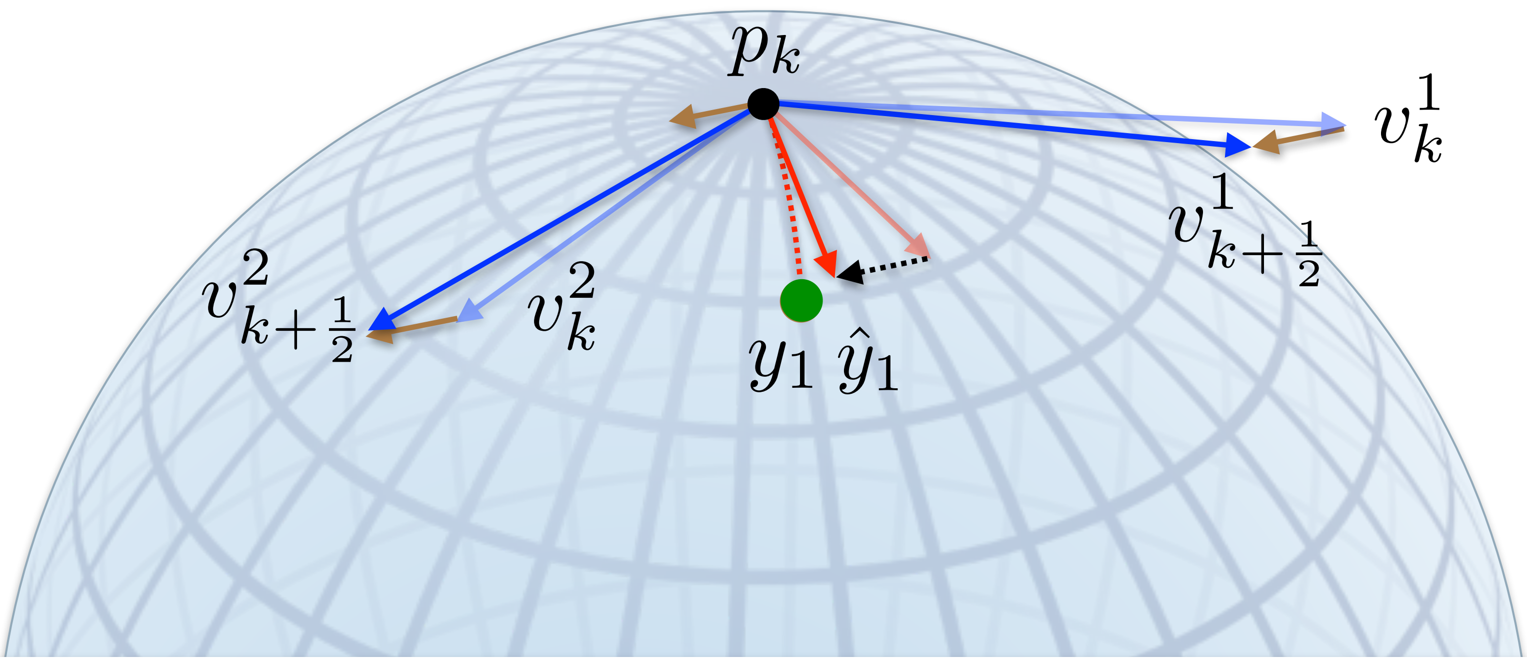
Parallel transported Error



Optimization (iterative method)

$$\mathrm{E}(p, \boldsymbol{V}) = \frac{1}{2} \sum_{i=1}^{N} \mathrm{d}(\mathrm{Exp}(p, \sum_j x_i^j v^j), y_i)^2$$

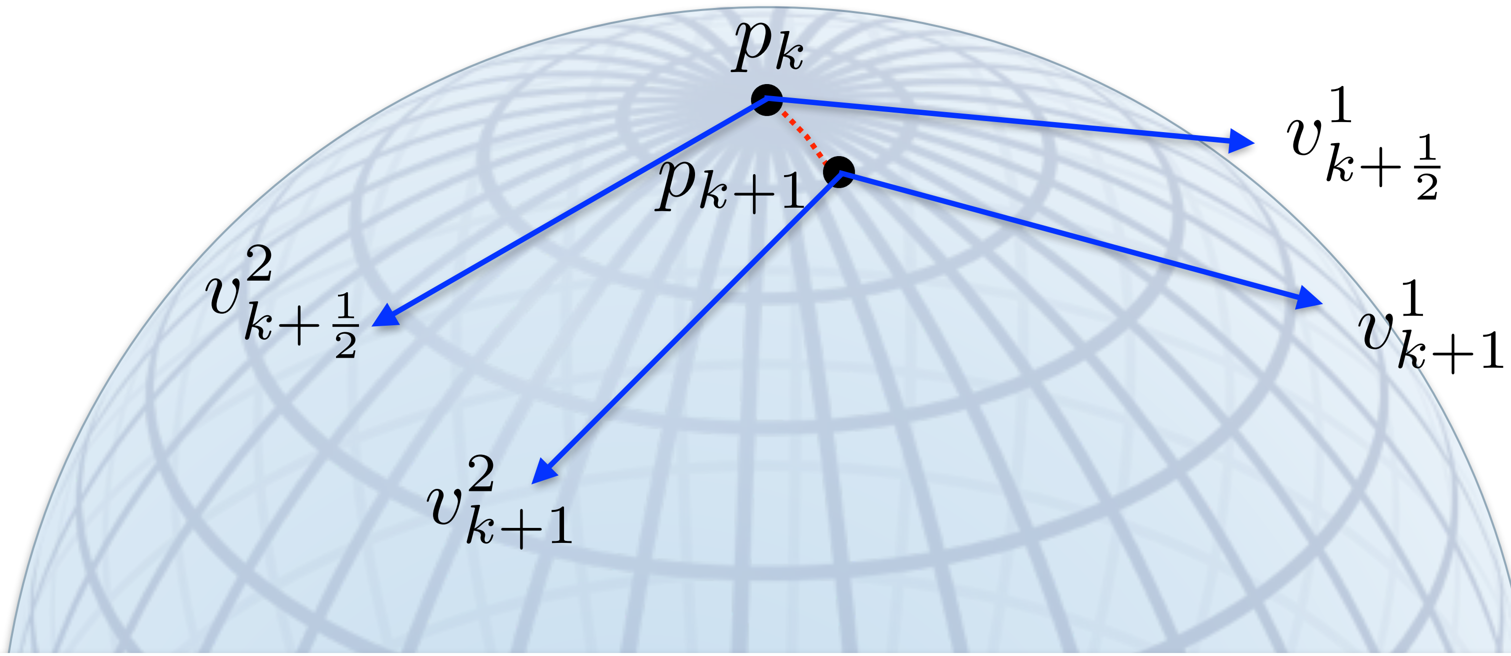
Step 2 : $$v_{k+1/2}^j = v_k^j + \alpha \sum_i x_i^j \Gamma_{\hat{y}_i \to p} \mathrm{Log}(\hat{y}_i, y_i)$$



Optimization (iterative method)

$$\mathrm{E}(p, \boldsymbol{V}) = \frac{1}{2} \sum_{i=1}^{N} \mathrm{d}(\mathrm{Exp}(p, \sum_j x_i^j v^j), y_i)^2$$

Step 3 : $$\boldsymbol{V}_{k+1} = \Gamma_{p_k \to p_{k+1}}(\boldsymbol{V}_{k+1/2})$$

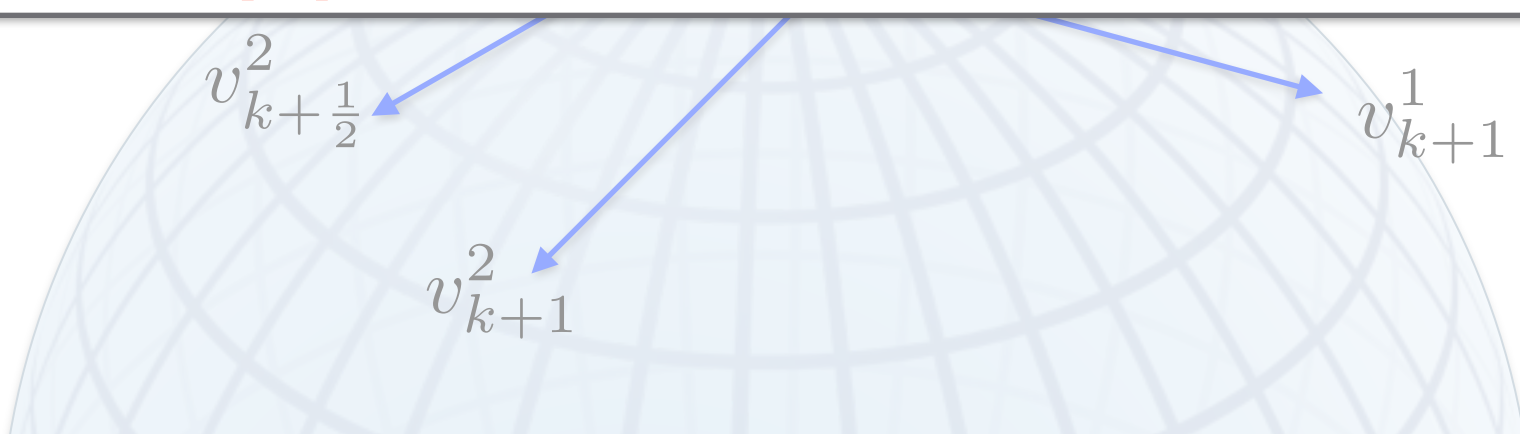


Optimization (iterative method)

$$\mathrm{E}(p, \boldsymbol{V}) = \frac{1}{2} \sum_{i=1}^{N} \mathrm{d}(\mathrm{Exp}(p, \sum_{j} x_i^j v^j), y_i)^2$$

Step 3 : $\boldsymbol{V}_{k+1} = \Gamma_{p_k \rightarrow p_{k+1}}(\boldsymbol{V}_{k+1/2})$

Must we solve for p or can we approximate it?

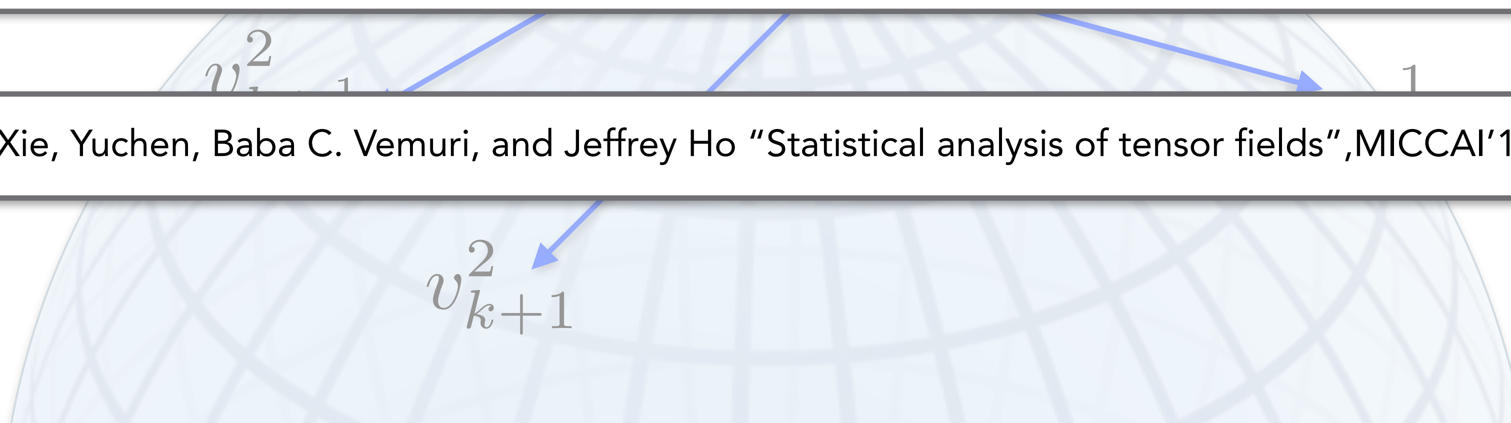


Optimization (iterative method)

$$\mathrm{E}(p, \boldsymbol{V}) = \frac{1}{2} \sum_{i=1}^{N} \mathrm{d}(\mathrm{Exp}(p, \sum_j x_i^j v^j), y_i)^2$$

Step 3 : $\boldsymbol{V}_{k+1} = \Gamma_{p_k \to p_{k+1}}(\boldsymbol{V}_{k+1/2})$

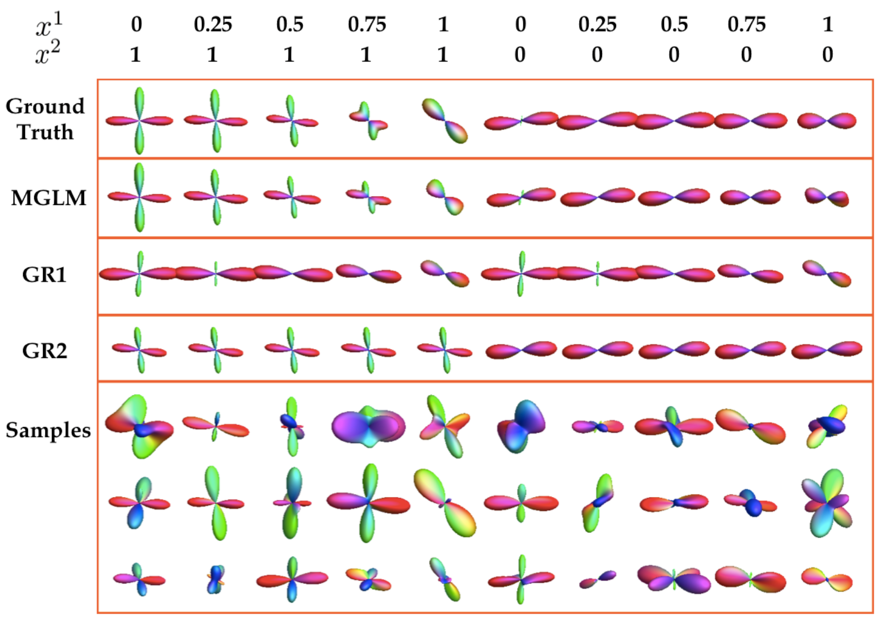
Must we solve for p or can we approximate it?



Xie, Yuchen, Baba C. Vemuri, and Jeffrey Ho “Statistical analysis of tensor fields”,MICCAI’10

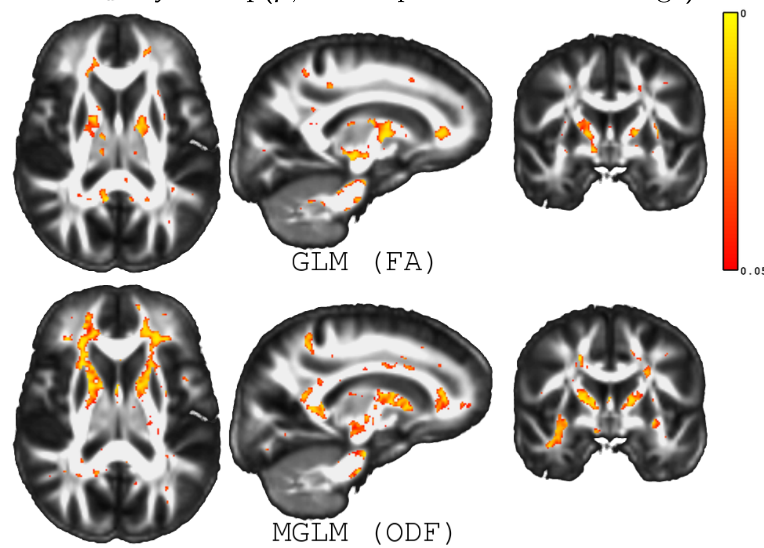
Experiments

Synthetic data



Neuroimaging data

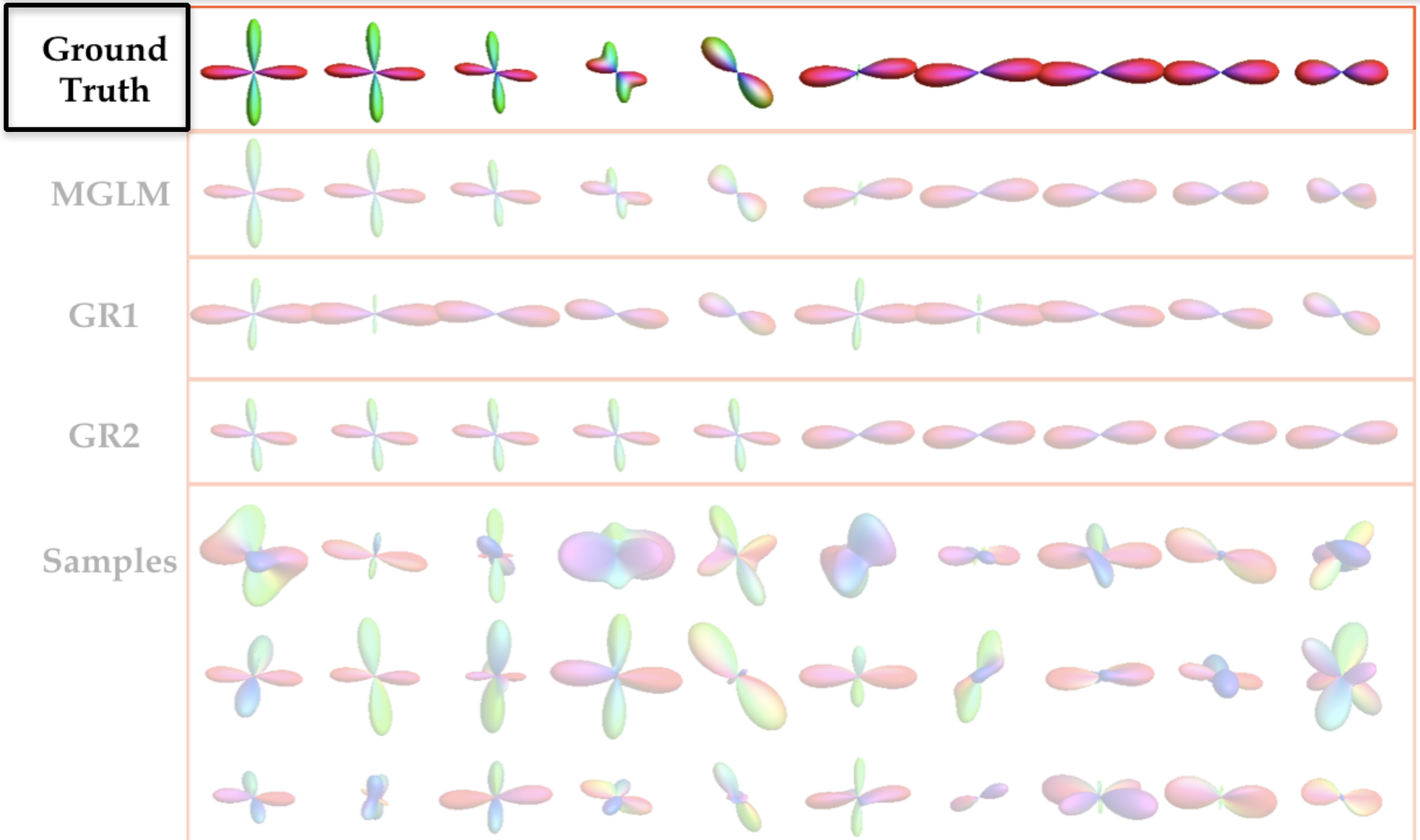
$\text{GLM}_{\text{Full}} : y = \text{Exp}(p, v^1\text{Group} + v^2\text{Gender} + v^3\text{Age})$



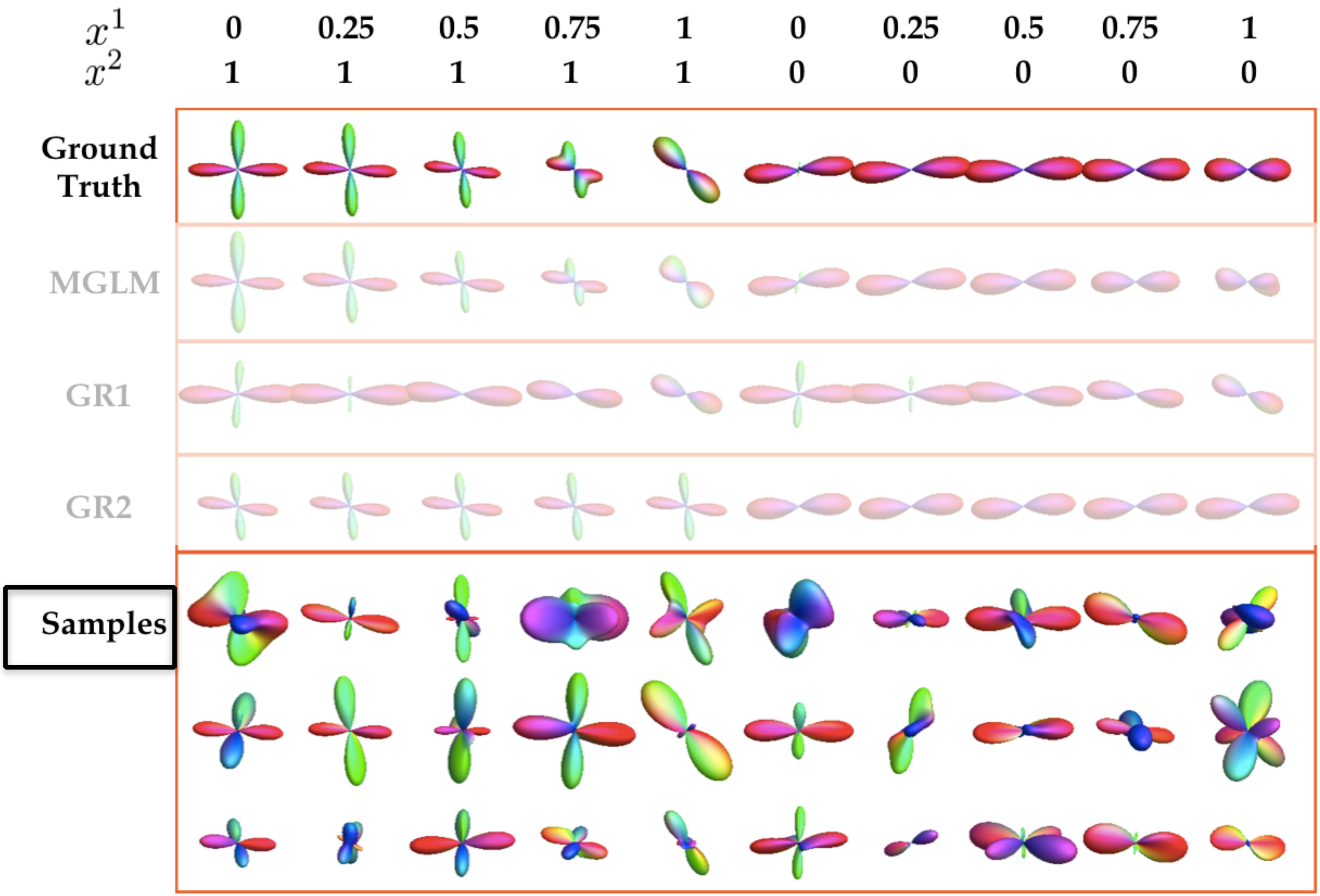
WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

Experiments (synthetic ODF)

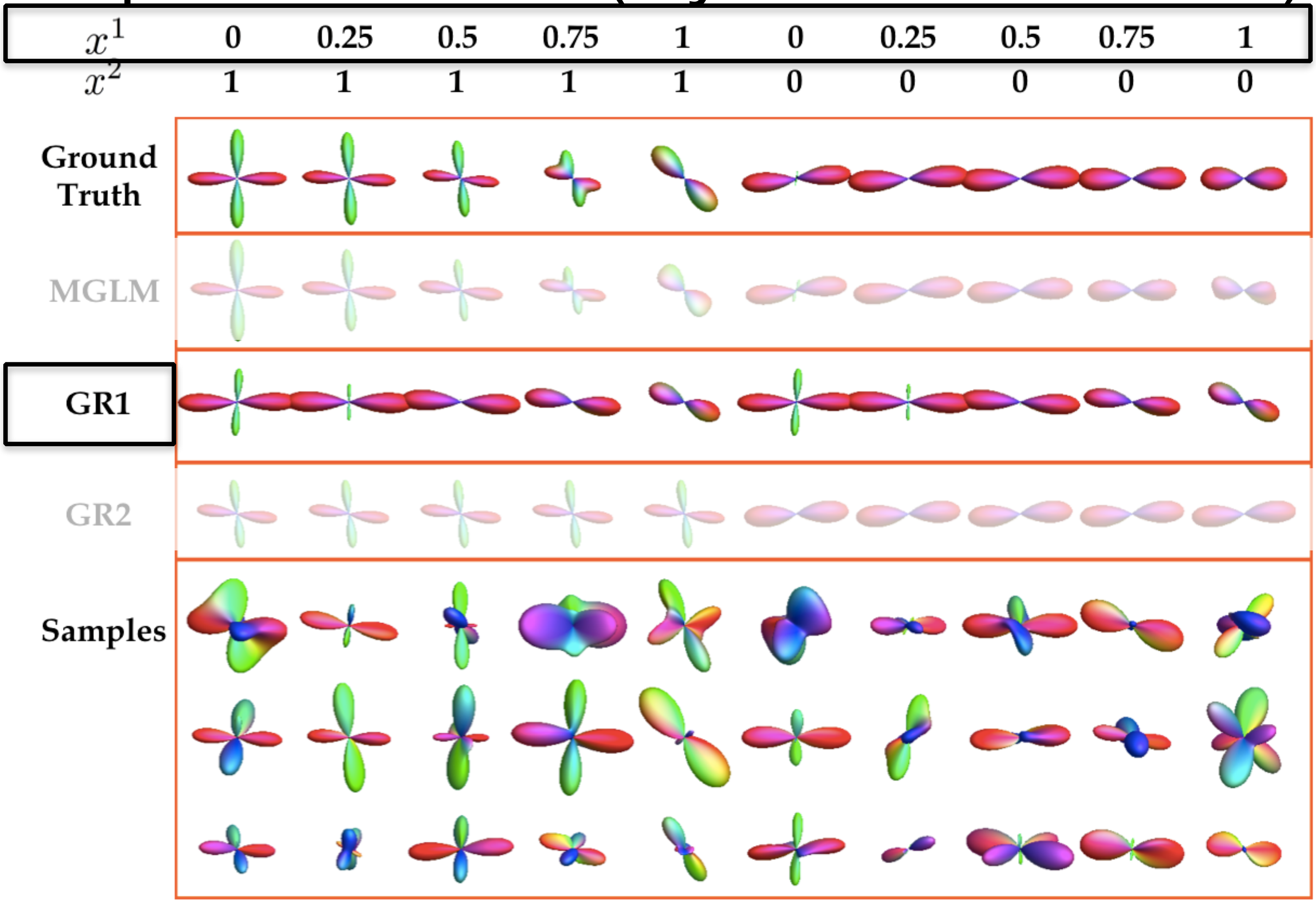
x^1	0	0.25	0.5	0.75	1	0	0.25	0.5	0.75	1
x^2	1	1	1	1	1	0	0	0	0	0



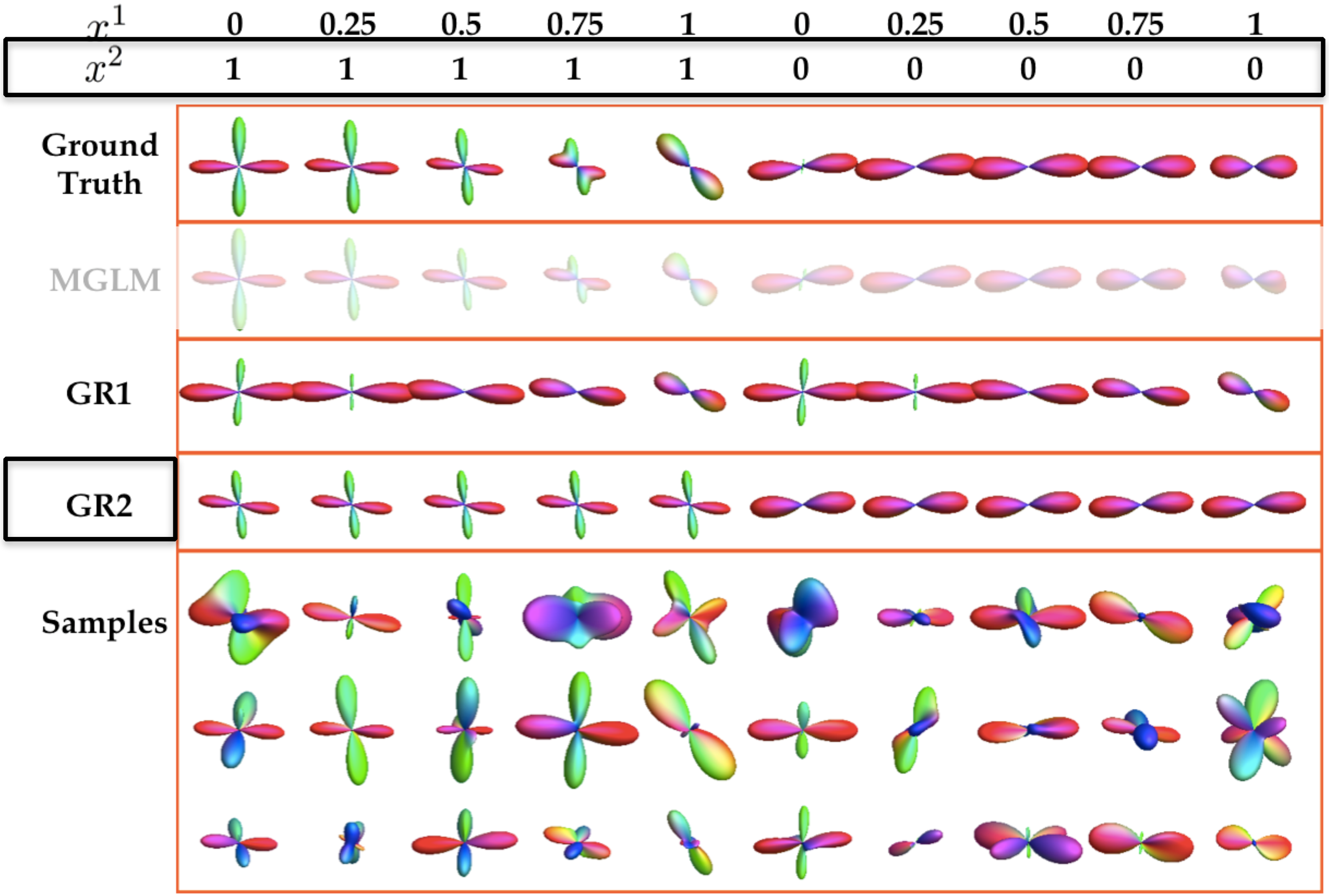
Experiments (synthetic ODF)



Experiments (synthetic ODF)

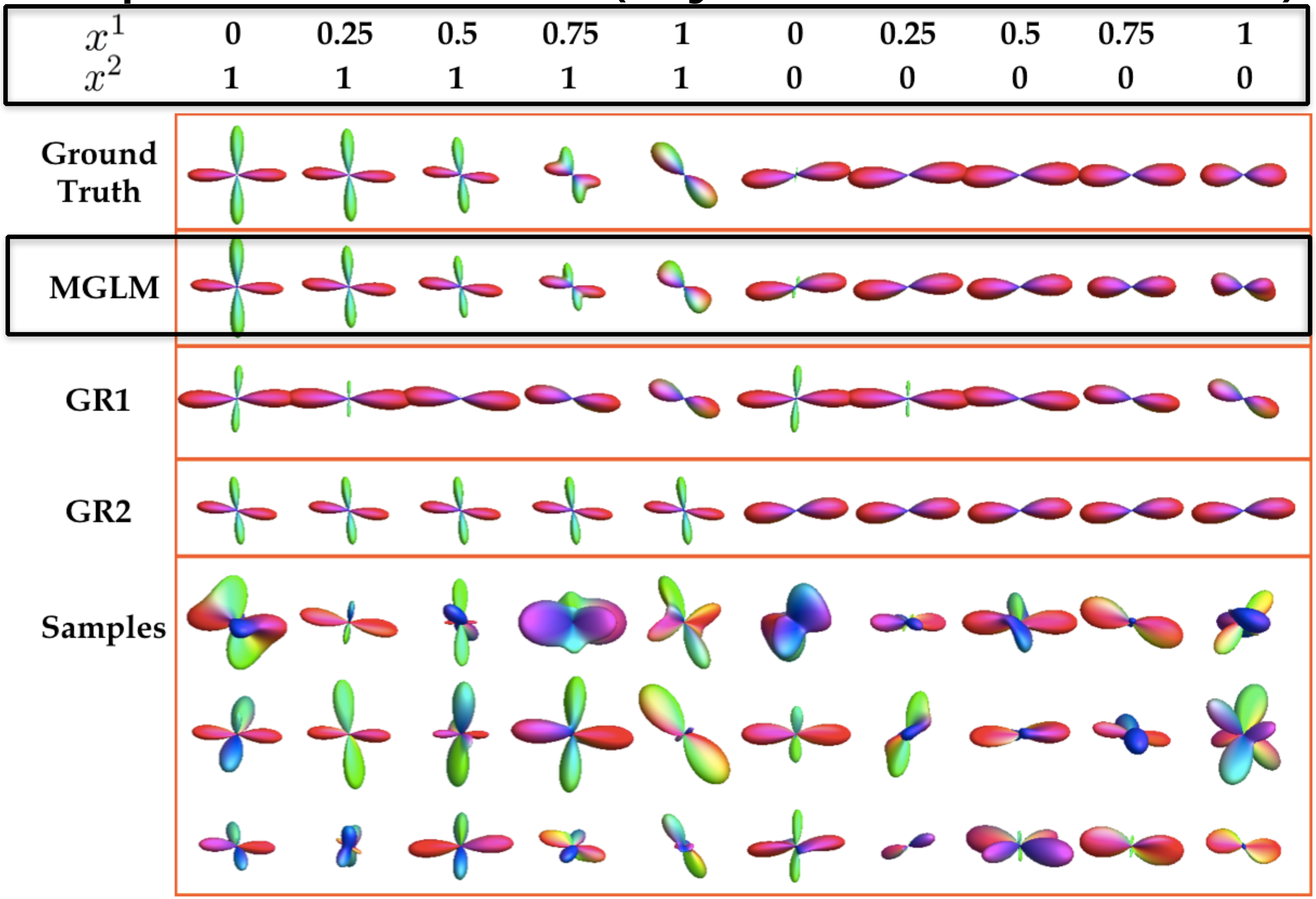


Experiments (synthetic ODF)



WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

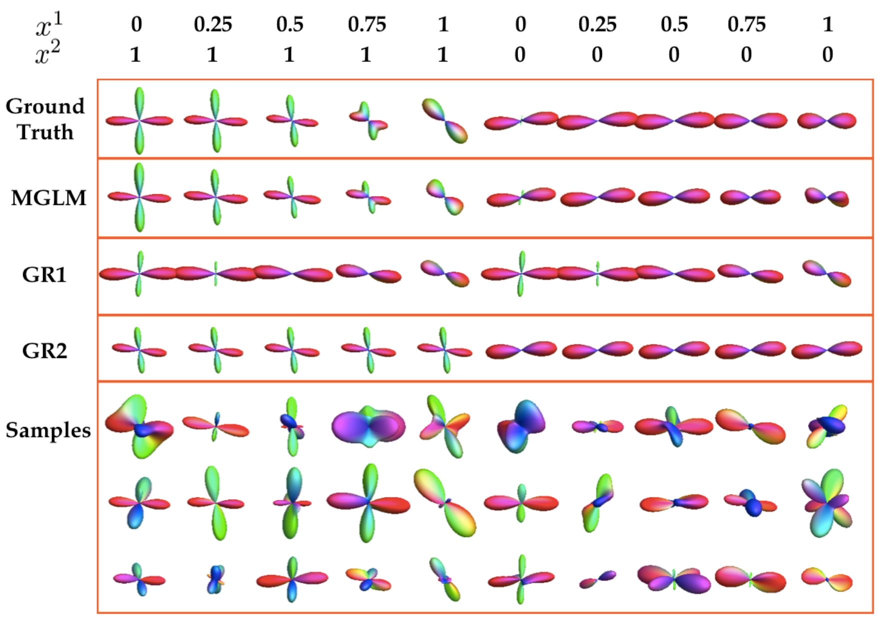
Experiments (synthetic ODF)



WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

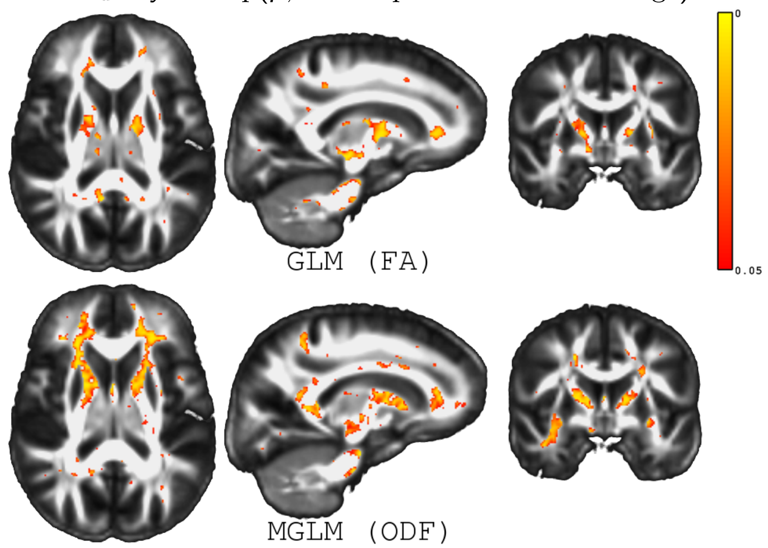
Experiments

Synthetic data



Neuroimaging data

$$\mathrm{GLM}_{\mathrm{Full}} : y = \mathrm{Exp}(p, v^1 \mathrm{Group} + v^2 \mathrm{Gender} + v^3 \mathrm{Age})$$



WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

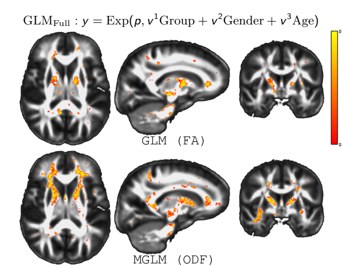
Experiments (neuroimaging study)

	LTM (ODF)		AD risk (DTI)	
Subjects	49		343	
Group	LTM	WLC	APOE4+	APOE4-
Gender	Female	Male	Female	Male
Age	28 - 65		43 - 75	

$$GLM_{Full} : y = \text{Exp}(p, v^1\text{Group} + v^2\text{Gender} + V^3\text{Age})$$

$$GLM_{Age} : y = \text{Exp}(p, v^2\text{Gender} + v^3\text{Age})$$

$$GLM_{Group} : y = \text{Exp}(p, v^1\text{Group} + v^2\text{Gender}$$

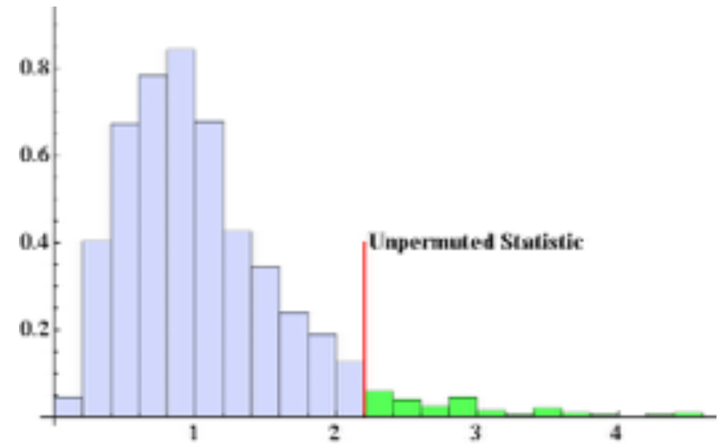


Experiments (neuroimaging study)

- *p*-values maps and histograms for effect of age and group computed from simulating the Null distribution of the F ratio statistic using 20,000 permutations.

- *F* ratio statistic is defined for a pair of nested GLMs as

$$F = \frac{\frac{\mathrm{RSS}_1 - \mathrm{RSS}_2}{p_2 - p_1}}{\frac{\mathrm{RSS}_2}{N - p_2}}$$

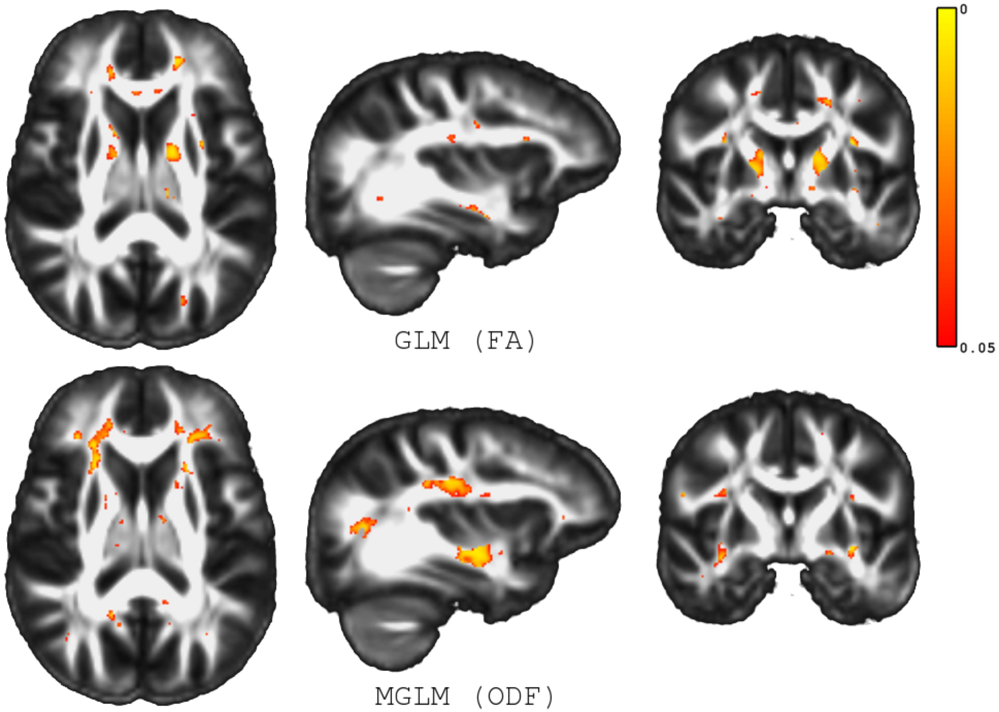


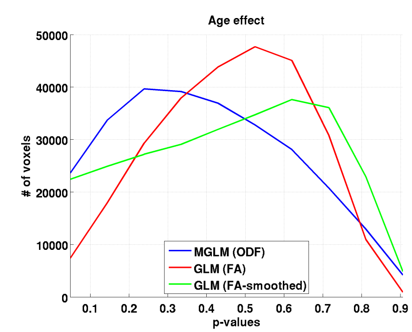
WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

Neuroimaging study 1

- Age effect in study 1 (long-term meditators, ODFs)

$$\mathrm{GLM}_{\mathrm{Full}} : y = \mathrm{Exp}(p, v^1 \mathrm{Group} + v^2 \mathrm{Gender} + v^3 \mathrm{Age})$$
$$\mathrm{GLM}_{\mathrm{Group}} : y = \mathrm{Exp}(p, v^1 \mathrm{Group} + v^2 \mathrm{Gender})$$

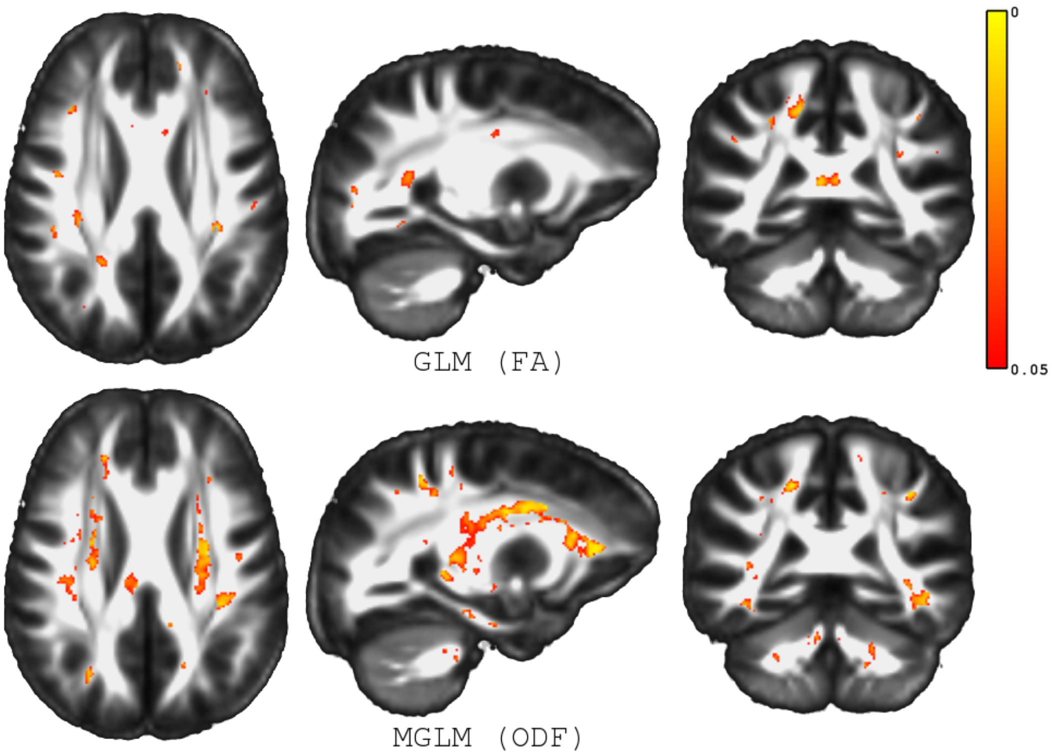


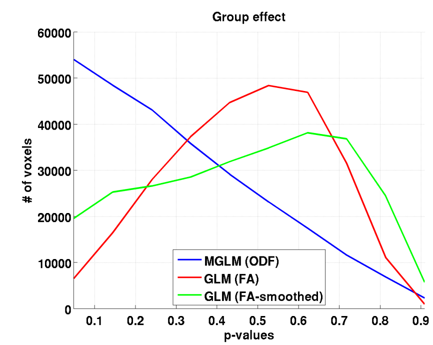


Neuroimaging study 1

- Group effect in study 1 (long-term meditators, ODFs)

$$\mathrm{GLM}_{\mathrm{Full}} : y = \mathrm{Exp}(p, v^1\mathrm{Group} + v^2\mathrm{Gender} + v^3\mathrm{Age})$$
$$\mathrm{GLM}_{\mathrm{Age}} : y = \mathrm{Exp}(p, v^2\mathrm{Gender} + v^3\mathrm{Age})$$





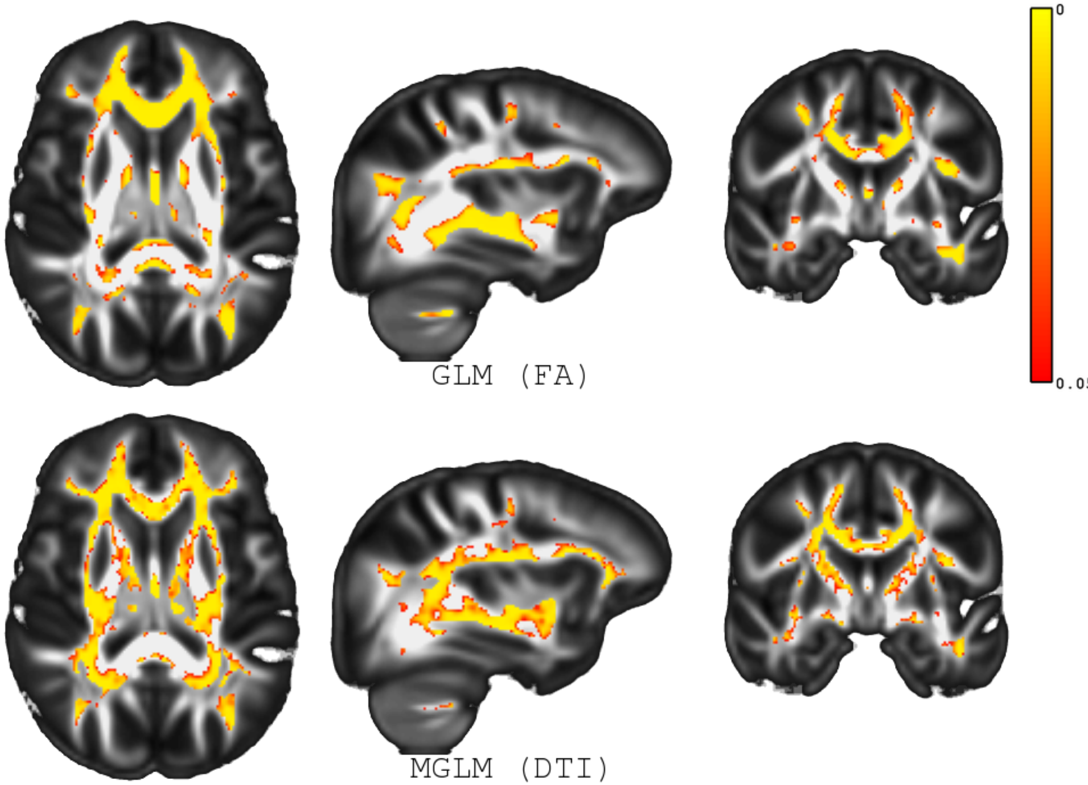
$\mathsf{Group} \in \{\mathsf{LTM}, \mathsf{WLC}\}.$

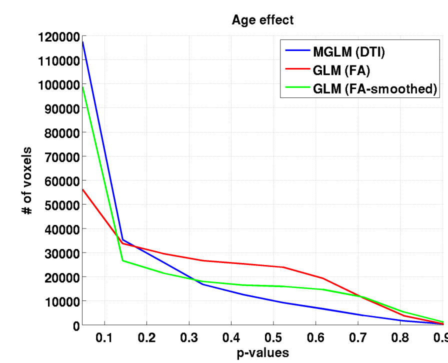
WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

Neuroimaging study 2

- Age model of study 2 (AD, DTI)

$$\mathrm{GLM}_{\mathrm{Full}} : y = \mathrm{Exp}(p, v^1\mathrm{Group} + v^2\mathrm{Gender} + v^3\mathrm{Age})$$
$$\mathrm{GLM}_{\mathrm{Group}} : y = \mathrm{Exp}(p, v^1\mathrm{Group} + v^2\mathrm{Gender})$$





WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

Conclusion

- Generalization of ***multivariate general linear model*** (MGLM) to Riemannian manifolds
- Especially useful when response is manifold valued and we want to control for one or more covariates. Here, the analysis obtains significantly improved statistical power
- Applicable to other manifold-valued statistical inference problems
- Code is available. See me at the posters!

Poster session ID: O-3C-384

Thank you

Research supported in part by
NSF CAREER RI 1252725
NIH R01 AG040396

Poster session ID: O-3C-384